08-21-06

IAP15 Rec'd PCT/PTO 17 AUG 2006

Express Mail Label No.: EV 532 353 144 US
Date of Deposit: August 17, 2006

Attorney Docket No. 27353-510-059

OIPE IAP14 AUG 17 2006 PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS :	Boutell, J.M. et al	CONFIRMATION NUMBER :	9992
SERIAL NUMBER :	10/527,603	EXAMINER :	Not yet assigned
INT'L FILING DATE :	September 16, 2003	INT'L APPLICATION NUMBER:	PCT/IB03/05258
FILING DATE :	March 15, 2005	ART UNIT :	Not yet assigned
FOR :	**ARRAYS AND METHODS**		

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL LETTER

Transmitted herewith for filing in the above-referenced patent application are the following documents:

1. Certified copy of PCT/GB2002/005499 and
2. Return postcard.

If the enclosed papers are considered incomplete, the Mail Room and/or the Application Branch is respectfully requested to contact the undersigned at (212) 935-3000, New York, New York. A duplicate copy of this transmittal letter is enclosed.

The Commissioner is authorized to charge any additional fees that may be due, or to credit any overpayment, to the undersigned's account, Deposit Account No. **50-0311**, Ref. No. **27353-510-059**, Customer Number: **35437**.

Respectfully submitted,

Dated: August 17, 2006

Michelle A. Klein

Ivor R. Elrifi, Reg. No. 39,529
Michelle A. Klein, Reg. No. 55,296
Attorneys/Agents for Applicants
c/o MINTZ, LEVIN, *et al.*
666 Third Avenue-24th Floor
New York, New York 10017
Telephone: (212) 983-3000
Telefax: (212) 983-3115

WORLD INTELLECTUAL PROPERTY ORGANIZATION

世界知识产权组织

ORGANIZACIÓN MUNDIAL DE LA PROPIEDAD INTELECTUAL

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE

المنظمة العالمية للملكية الفكرية

ВСЕМИРНАЯ ОРГАНИЗАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

CERTIFICATION

It is hereby certified that the attached copy is a true copy of the record copy of International Application No. GB2002/005499, filed with the United Kingdom Patent Office as receiving Office on 5 December 2002 (05/12/2002) and received by the International Bureau on 21 January 2003 (21/01/2003), including any pages containing corrections and/or rectifications transmitted by the competent Authority to, and received by, the International Bureau before the completion of the technical preparations for international publication.

By: The International Bureau

Fabienne Gateau
Senior PCT Assistant
PCT Legal Affairs Section
PCT Legal Division

Date: 11 April 2006 (11/04/2006)

PCT

REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

* title changed See ISR

For receiving Office use only
PCT/GB 2002 / 0 0 5 4 9 9
International Application No.
International Filing Date 5 DECEMBER 2002
Name of receiving Office and "PCT International Application" — United Kingdom Patent Office PCT International Application

Applicant's or agent's file reference (if desired) (12 characters maximum) PWC/P33293~~US~~ WO

Box No. I TITLE OF INVENTION

[ASSAYS] *

see 09

Box No. II APPLICANT ☐ This person is also inventor

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

Sense Proteomic Limited
Babraham Hall *
Babraham
Cambridge CB2 4AT
~~[UK]~~ ▲▲ GB ▲

* see # 11

▲▲ RO/GB Deleted ▲ RO/GB Added

Telephone No.

Facsimile No.

Teleprinter No.

Applicant's registration No. with the Office

State *(that is, country)* of nationality: GB

State *(that is, country)* of residence: GB

This person is applicant for the purposes of: ☐ all designated States ☒ all designated States except the United States of America ☐ the United States of America only ☐ the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

BOUTELL, Jonathan Mark
Sense Proteomic Limited
Babraham Hall
Babraham
Cambridge, CB2 4AT, ~~[UK]~~ ▲▲ GB ▲

This person is:
☐ applicant only
☒ applicant and inventor
☐ inventor only *(If this check-box is marked, do not fill in below.)*

Applicant's registration No. with the Office

State *(that is, country)* of nationality: GB

State *(that is, country)* of residence: GB

This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America ☒ the United States of America only ☐ the States indicated in the Supplemental Box

☒ Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: ☒ agent ☐ common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

CHAPMAN, Paul William
Kilburn & Strode
20 Red Lion Street
London WC1R 4PJ
United Kingdom

Telephone No. 020 7539 4200

Facsimile No. 020 7539 4299

Teleprinter No.

Agent's registration No. with the Office

☐ **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

PCT/GB 2002/005499

Sheet No. ...2...

Continuation of Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

If none of the following sub-boxes is used, this sheet should not be included in the request.

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*	This person is:
GODBER, Benjamin Leslie James Sense Proteomic Limited Babraham Hall Babraham Cambridgeshire, CB2 4AT, ~~UK~~▲▲ GB▲	☐ applicant only ☒ applicant and inventor ☐ inventor only *(If this check-box is marked, do not fill in below.)* Applicant's registration No. with the Office
State *(that is, country)* of nationality: GB	State *(that is, country)* of residence: GB
This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America	☒ the United States of America only ☐ the States indicated in the Supplemental Box

▲ RO/GB Added

▲▲ RO/GB Deleted

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*	This person is:
HART, Darren James Sense Proteomic Limited Babraham Hall Babraham Cambridgeshire, CB2 4AT, ~~UK~~▲▲ GB▲	☐ applicant only ☒ applicant and inventor ☐ inventor only *(If this check-box is marked, do not fill in below.)* Applicant's registration No. with the Office
State *(that is, country)* of nationality: GB	State *(that is, country)* of residence: GB
This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America	☒ the United States of America only ☐ the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*	This person is:
BLACKBURN, Jonathan Michael Sense Proteomic Limited Babraham Hall Babraham Cambridgeshire, CB2 4AT, ~~UK~~▲▲ GB▲ *	☐ applicant only ☒ applicant and inventor ☐ inventor only *(If this check-box is marked, do not fill in below.)* Applicant's registration No. with the Office
State *(that is, country)* of nationality: GB	State *(that is, country)* of residence: GB
This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America	☒ the United States of America only ☐ the States indicated in the Supplemental Box

*see #11

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*	This person is:
	☐ applicant only ☐ applicant and inventor ☐ inventor only *(If this check-box is marked, do not fill in below.)* Applicant's registration No. with the Office
State *(that is, country)* of nationality:	State *(that is, country)* of residence:
This person is applicant for the purposes of: ☐ all designated States ☐ all designated States except the United States of America	☐ the United States of America only ☐ the States indicated in the Supplemental Box

☐ Further applicants and/or (further) inventors are indicated on another continuation sheet.

Box No. V DESIGNATION OF STATES *Mark the applicable check-boxes below; at least one must be marked.*

The following designations are hereby made under Rule 4.9(a):

Regional Patent

☒ AP **ARIPO Patent:** GH Ghana, GM Gambia, KE Kenya, LS Lesotho, MW Malawi, MZ Mozambique, SD Sudan, SL Sierra Leone, SZ Swaziland, TZ United Republic of Tanzania, UG Uganda, ZM Zambia, ZW Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* ....

☒ EA **Eurasian Patent:** AM Armenia, AZ Azerbaijan, BY Belarus, KG Kyrgyzstan, KZ Kazakhstan, MD Republic of Moldova, RU Russian Federation, TJ Tajikistan, TM Turkmenistan, and any other State which is a Contracting State of the Eurasian Patent Convention and of the PCT

☒ EP **European Patent:** AT Austria, BE Belgium, BG Bulgaria, CH & LI Switzerland and Liechtenstein, CY Cyprus, CZ Czech Republic, DE Germany, DK Denmark, EE Estonia, ES Spain, FI Finland, FR France, GB United Kingdom, GR Greece, IE Ireland, IT Italy, LU Luxembourg, MC Monaco, NL Netherlands, PT Portugal, SE Sweden, SK Slovakia, TR Turkey, and any other State which is a Contracting State of the European Patent Convention and of the PCT

☒ OA **OAPI Patent:** BF Burkina Faso, BJ Benin, CF Central African Republic, CG Congo, CI Côte d'Ivoire, CM Cameroon, GA Gabon, GN Guinea, GQ Equatorial Guinea, GW Guinea-Bissau, ML Mali, MR Mauritania, NE Niger, SN Senegal, TD Chad, TG Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* ....

National Patent *(if other kind of protection or treatment desired, specify on dotted line)*:

☒ AE United Arab Emirates ....
☒ AG Antigua and Barbuda
☒ AL Albania ....
☒ AM Armenia ....
☒ AT Austria ....
☒ AU Australia ....
☒ AZ Azerbaijan ....
☒ BA Bosnia and Herzegovina ....
☒ BB Barbados
☒ BG Bulgaria ....
☒ BR Brazil ....
☒ BY Belarus ....
☒ BZ Belize ....
☒ CA Canada
☒ CH & LI Switzerland and Liechtenstein
☒ CN China ....
☒ CO Colombia
☒ CR Costa Rica ....
☒ CU Cuba ....
☒ CZ Czech Republic ....
☒ DE Germany ....
☒ DK Denmark ....
☒ DM Dominica
☒ DZ Algeria ....
☒ EC Ecuador ....
☒ EE Estonia ....
☒ ES Spain ....
☒ FI Finland ....
☒ GB United Kingdom
☒ GD Grenada
☒ GE Georgia ....
☒ GH Ghana ....
☒ GM Gambia
☒ HR Croatia ....
☒ HU Hungary ....
☒ ID Indonesia
☒ IL Israel ....
☒ IN India ....
☒ IS Iceland
☒ JP Japan ....
☒ KE Kenya ....
☒ KG Kyrgyzstan ....
☒ KP Democratic People's Republic of Korea ....
☒ KR Republic of Korea ....
☒ KZ Kazakhstan ....
☒ LC Saint Lucia
☒ LK Sri Lanka
☒ LR Liberia
☒ LS Lesotho ....
☒ LT Lithuania
☒ LU Luxembourg
☒ LV Latvia
☒ MA Morocco ....
☒ MD Republic of Moldova ....
....
☒ MG Madagascar ....
☒ MK The former Yugoslav Republic of Macedonia ....
☒ MN Mongolia
☒ MW Malawi ....
☒ MX Mexico ....
☒ MZ Mozambique ....
☒ NO Norway
☒ NZ New Zealand ....
☒ OM Oman
☒ PH Philippines ....
☒ PL Poland ....
☒ PT Portugal ....
☒ RO Romania
☒ RU Russian Federation ....
....
☒ SD Sudan
☒ SE Sweden
☒ SG Singapore
☒ SI Slovenia ....
☒ SK Slovakia ....
☒ SL Sierra Leone ....
☒ TJ Tajikistan ....
☒ TM Turkmenistan ....
☒ TN Tunisia
☒ TR Turkey ....
☒ TT Trinidad and Tobago ....
....
☒ TZ United Republic of Tanzania
☒ UA Ukraine ....
☒ UG Uganda ....
☒ US United States of America ....
....
☒ UZ Uzbekistan ....
☒ VN Viet Nam ....
☒ YU Yugoslavia ....
☒ ZA South Africa ....
☒ ZM Zambia
☒ ZW Zimbabwe ....

Check-boxes below reserved for designating States which have become party to the PCT after issuance of this sheet:

☒ St Vincent & Grenadines .... ☐ .... ☐ ....
☐ .... ☐ .... ☐ ....

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. *(Confirmation (including fees) must reach the receiving Office within the 15-month time limit.)*

Supplemental Box *If the Supplemental Box is not used, this sheet should not be included in the request.*

1. *If, in any of the Boxes, except Boxes Nos. VIII(i) to (v) for which a special continuation box is provided, **the space is insufficient** to furnish all the information: in such case, write "Continuation of Box No...." (indicate the number of the Box) and furnish the information in the same manner as required according to the captions of the Box in which the space was insufficient, in particular:*

(i) ***if more than two persons are to be indicated as applicants and/or inventors** and no "continuation sheet" is available: in such case, write "Continuation of Box No. III" and indicate for each additional person the same type of information as required in Box No. III. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below;*

(ii) *if, in Box No. II or in any of the sub-boxes of Box No. III, the indication **"the States indicated in the Supplemental Box"** is checked: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the applicant(s) involved and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is applicant;*

(iii) *if, in Box No. II or in any of the sub-boxes of Box No. III, **the inventor or the inventor/applicant is not inventor for the purposes of all designated States or for the purposes of the United States of America**: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the inventor(s) and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is inventor;*

(iv) *if, in addition to the agent(s) indicated in Box No. IV, there are **further agents**: in such case, write "Continuation of Box No. IV" and indicate for each further agent the same type of information as required in Box No. IV;*

(v) *if, in Box No. V, the name of any State (or OAPI) is accompanied by the indication **"patent of addition,"** or **"certificate of addition,"** or if, in Box No. V, the name of the United States of America is accompanied by an indication **"continuation"** or **"continuation-in-part"**: in such case, write "Continuation of Box No. V" and the name of each State involved (or OAPI), and after the name of each such State (or OAPI), the number of the parent title or parent application and the date of grant of the parent title or filing of the parent application;*

(vi) *if, in Box No. VI, there are **more than five earlier applications whose priority is claimed**: in such case, write "Continuation of Box No. VI" and indicate for each additional earlier application the same type of information as required in Box No. VI.*

2. *If, with regard to the **precautionary designation statement** contained in Box No. V, the applicant wishes to exclude any State(s) from the scope of that statement: in such case, write "Designation(s) excluded from precautionary designation statement" and indicate the name or two-letter code of each State so excluded.*

Additional Representatives

Ashmead,	Richard John
Jennings,	Nigel Robin
Rees,	David Christopher
Maggs,	Michael Norman
Hale,	Peter
Miller,	James Lionel Woolverton
Roberts,	Gwilym Vaughan
Cornish,	Kristina Victoria Joy
Gold,	Tibor Zoltan
Hedley,	Nicholas James Matthew
Bassil,	Nicholas Charles
Lee,	Nicholas John
Copsey,	Timothy Graham
Hibbert,	Juliet Jane Grace
Addison,	Ann Bridget
Ford,	Timothy

All of: Kilburn & Strode
20 Red Lion Street
London WC1R 4PJ
United Kingdom

Box No. VI PRIORITY CLAIM

The priority of the following earlier application(s) is hereby claimed:

Filing date of earlier application *(day/month/year)*	Number of earlier application	Where earlier application is: national application: country or Member of WTO	regional application:* regional Office	international application: receiving Office
item (1) 5/12/01 5 DECEMBER 2001 ▲	60/335,806	US		
item (2) 16/09/02 16 SEPTEMBER 2002 ▲	60/410,815	US		
item (3)				
item (4)				
item (5)				

▲ RO/GB Added

☐ Further priority claims are indicated in the Supplemental Box.

The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of this international application is the receiving Office)* identified above as:

☐ all items ☒ item (1) ☒ item (2) ☐ item (3) ☐ item (4) ☐ item (5) ☐ other, see Supplemental Box

* *Where the earlier application is an ARIPO application, indicate at least one country party to the Paris Convention for the Protection of Industrial Property or one Member of the World Trade Organization for which that earlier application was filed (Rule 4.10(b)(ii)):* ....

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used)*:

ISA / ....

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority)*:

Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII DECLARATIONS

The following **declarations** are contained in Boxes Nos. VIII (i) to (v) *(mark the applicable check-boxes below and indicate in the right column the number of each type of declaration)*:

			Number of declarations
☐	Box No. VIII (i)	Declaration as to the identity of the inventor	:
☐	Box No. VIII (ii)	Declaration as to the applicant's entitlement, as at the international filing date, to apply for and be granted a patent	:
☐	Box No. VIII (iii)	Declaration as to the applicant's entitlement, as at the international filing date, to claim the priority of the earlier application	:
☐	Box No. VIII (iv)	Declaration of inventorship (only for the purposes of the designation of the United States of America)	:
☐	Box No. VIII (v)	Declaration as to non-prejudicial disclosures or exceptions to lack of novelty	:

PCT/GB 2002 / 0 0 5 4 9 9

Sheet No. . . .6. . .

Box No. IX CHECK LIST; LANGUAGE OF FILING

This international application **contains**:

(a) the following number of **sheets in paper form**:

request (including declaration sheets)	:	6
description (excluding sequence listing part)	:	57
claims	:	4
abstract	:	1
drawings	:	27
Sub-total number of sheets	:	95
sequence listing part of description *(actual number of sheets if filed in paper form, whether or not also filed in computer readable form; see (b) below)*	:	
Total number of sheets	:	95

(b) sequence listing part of description filed in **computer readable form**

(i) ☐ only (under Section 801(a)(i))

(ii) ☐ in addition to being filed in paper form (under Section 801(a)(ii))

Type and number of carriers (diskette, CD-ROM, CD-R or other) on which the sequence listing part is contained *(additional copies to be indicated under item 9(ii), in right column)*: . . . .

This international application is **accompanied by** the following item(s) *(mark the applicable check-boxes below and indicate in right column the number of each item)*: — Number of items

1. ☐ fee calculation sheet :
2. ☐ original separate power of attorney :
3. ☐ original general power of attorney :
4. ☐ copy of general power of attorney; reference number, if any: . . . . :
5. ☐ statement explaining lack of signature :
6. ☐ priority document(s) identified in Box No. VI as item(s): . . . . :
7. ☐ translation of international application into *(language)*: . . . . :
8. ☐ separate indications concerning deposited microorganism or other biological material :
9. ☐ sequence listing in computer readable form (indicate also type and number of carriers (diskette, CD-ROM, CD-R or other))
   (i) ☐ copy submitted for the purposes of international search under Rule 13*ter* only (and not as part of the international application) :
   (ii) ☐ *(only where check-box (b)(i) or (b)(ii) is marked in left column)* additional copies including, where applicable, the copy for the purposes of international search under Rule 13*ter* :
   (iii) ☐ together with relevant statement as to the identity of the copy or copies with the sequence listing part mentioned in left column :
10. ☐ other *(specify)*: . . . . :

Figure of the drawings which should accompany the abstract:

Language of filing of the international application: English

Box No. X SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

5 December 2002

CHAPMAN, Paul William
Agent for the Applicants

For receiving Office use only

1. Date of actual receipt of the purported international application: 05 DECEMBER 2002

2. Drawings:
☒ received:
☐ not received:

3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:

4. Date of timely receipt of the required corrections under PCT Article 11(2):

5. International Searching Authority (if two or more are competent): ISA /

6. ☐ Transmittal of search copy delayed until search fee is paid

For International Bureau use only

Date of receipt of the record copy by the International Bureau: 21 JAN 2003

Form PCT/RO/101 (last sheet) (March 2001; reprint July 2002) *See Notes to the request form*

ARRAYS

Single nucleotide polymorphisms (SNPs) are single base differences between the DNA of organisms. They underlie much of the genetic component of phenotypic variation between individuals with the exception of identical siblings and clones. Since this variation includes characteristics such as predisposition to disease, age of onset, severity of disease and response to treatment, the identification and cataloguing of SNPs will lead to 'genetic medicine' [Chakravarti, A. *Nature* 409 822-823 (2001)]. Disciplines such as pharmacogenomics are aiming to establish correlations between SNPs and response to drug treatment in order to tailor therapeutic programmes to the individual person. More broadly, the role of particular SNPs in conditions such as sickle cell anaemia and Alzheimer's disease, and issues such as HIV resistance and transplant rejection, are well appreciated. However, correlations between SNPs and their phenotypes are usually derived from statistical analyses of population data and little attempt is made to elucidate the molecular mechanism of the observed phenotypic variation. Until the advent of high-throughput sequencing projects aimed at determining the complete sequence of the human genome [The International Human Genome Mapping Consortium *Nature* 409 860-921 (2001); Venter, J.C. *Science* 291 1304-1351 (2001)], only a few thousand SNPs had been identified. More recently 1.42 million SNPs were catalogued by a consortium of researchers in a paper accompanying the human sequence [The International SNP Map Working Group *Nature* 409 928-933 (2001)] of which 60,000 were present within genes ('coding' SNPs). Coding SNPs can be further classified according to whether or not they alter the amino acid sequence of the protein and where changes do occur, protein function may be affected resulting in phenotypic variation. Thus there is an

unmet need for apparatus and methodology capable of rapidly determining the phenotypes of this large volume of variant sequences.

The Inventors herein describe protein arrays and their use to assay, in a parallel fashion, the protein products of highly homologous or related DNA coding sequences.

By highly homologous or related it is meant those DNA coding sequences which share a common sequence and which differ only by one or more naturally occurring mutations such as single nucleotide polymorphisms, deletions or insertions, or those sequences which are considered to be haplotypes (a haplotype being a combination of variations or mutations on a chromosome, usually within the context of a particular gene). Such highly homologous or related DNA coding sequences are generally naturally occurring variants of the same gene.

Arrays according to the invention have multiple for example, two or more, individual proteins deposited in a spatially defined pattern on a surface in a form whereby the properties, for example the activity or function of the proteins can be investigated or assayed in parallel by interrogation of the array.

Protein arrays according to the invention and their use to assay the phenotypic changes in protein function resulting from mutations (for example, coding SNPs – i.e. those SNP mutations that still give rise to an expressed protein) differ completely to, and have advantages over, existing DNA based technologies for SNP and other mutational analyses [reviewed in Shi, M.M *Clin Chem* 47 164-72 (2001)]. These latter technologies include high-throughput sequencing and

electrophoretic methods for identifying new SNPs, or diagnostic technologies such as high density oligonucleotide arrays [e.g. Lindblad-Toh, K. *Nat Genet* 24 381-6 (2000)] or high-throughput, short-read sequencing techniques which permit profiling of an individuals gene of interest against known SNPs [e.g. Buetow, K.H. *Proc Natl Acad Sci USA* 98 581-4 (2001)]. Importantly, and in contrast to the invention described herein, the phenotypic effects of a polymorphism remain unknown when only analysed at the DNA level.

Indeed, the effects of coding SNPs on the proteins they encode are, with relatively few exceptions, uncharacterised. Examples of proteins with many catalogued SNPs but little functional data on the effect of these SNPs include p53, p10 (both cancer related) and the cytochrome P450s (drug metabolism). There are currently few if any methods capable of investigating the functionalities of SNP-encoded proteins with sufficiently high throughput required to handle the large volume of SNP data being generated. Bioinformatics, or computer modelling is possible, especially if a crystal structure is available, but the hypotheses generated still need to be verified experimentally (i.e. through biochemical assay). Frequently though, the role of the mutation remains unclear after bioinformatic or computer-based analysis. Therefore, protein arrays as provided by the invention offer the most powerful route to functional analysis of SNPs.

It would be possible to individually assay proteins derived from related DNA molecules, for example differing by one or more single nucleotide polymorphisms, in a test tube format, however the serial nature of this work and the large sample volumes involved make this approach cumbersome and unattractive. By arraying out the related proteins in a microtiter plate or on a

microscope slide, many different proteins (hundreds or thousands) can be assayed simultaneously using only small sample volumes (few microlitres only in the case of microarrays) thus making functional analysis of, for example, SNPs economically feasible. All proteins can be assayed together in the same experiment which reduces sources of error due to differential handling of materials. Additionally, tethering the proteins directly to a solid support facilitates binding assays which require unbound ligands to be washed away prior to measuring bound concentrations, a feature not available in solution based or single phase liquid assays.

Specific advantages over apparatus and methods currently known in the art provided by the arrays of the present invention are:

- massively parallel analysis of closely related proteins, for example those derived from coding SNPs, for encoded function
- sensitivity of analysis at least comparable to existing methods, if not better
- enables quantitative, comparative functional analysis in a manner not previously possible
- compatible with protein: protein, protein: nucleic acid, protein: ligand, or protein: small molecule interactions and post-translational modifications *in situ* "on-chip"
- parallel protein arrays according to the invention are spotting density independent
- microarray format enables analysis to be carried out using small volumes of potentially expensive ligands

- information provided by parallel protein arrays according to the invention will be extremely valuable for drug discovery, pharmacogenomics and diagnostics fields
- other useful parallel protein arrays may include proteins derived from non-natural (synthetic) mutations of a DNA sequence of interest. Such arrays can be used to investigate interactions between the variant protein thus produced and other proteins, nucleic acid molecules and other molecules, for example ligands or candidate/test small molecules. Suitable methods of carrying out such mutagenesis are described in Current Protocols in Molecular Biology, Volume 1, Chapter 8, Edited by Ausubel, FM, Brent, R, Kingston, RE, Moore, DD, Siedman, JG, Smith, JA, and Struhl, K.

Thus in one aspect, the invention provides a protein array comprising a surface upon which are deposited at spatially defined locations at least two protein moieties characterised in that said protein moieties are those of naturally occurring variants of a DNA sequence of interest.

A protein array as defined herein is a spatially defined arrangement of protein moieties in a pattern on a surface. Preferably the protein moieties are attached to the surface either directly or indirectly. The attachment can be non-specific (e.g. by physical absorption onto the surface or by formation of a non-specific covalent interaction). In a preferred embodiment the protein moieties are attached to the surface through a common marker moiety appended to each protein moiety. In another preferred embodiment, the protein moieties can be incorporated into a vesicle or liposome which is tethered to the surface.

A surface as defined herein is a flat or contoured area that may or may not be coated/derivatised by chemical treatment. For example, the area can be :
a glass slide,
one or more beads, for example a magnetised, derivatised and/or labelled bead as known in the art,
a polypropylene or polystyrene slide,
a polypropylene or polystyrene multi-well plate,
a gold, silica or metal object,
a membrane made of nitrocellulose, PVDF, nylon or phosphocellulose

Where a bead is used, individual proteins, pairs of proteins or pools of variant proteins (e.g., for "shotgun screening" - to initially identify groups of proteins in which a protein of interest may exist; such groups are then separated and further investigated (analogous to pooling methods known in the art of combinatorial chemistry)) may be attached to an individual bead to provide the spatial definition or separation of the array. The beads may then be assayed separately, but in parallel, in a compartmentalised way, for example in the wells of a microtitre plate or in separate test tubes.

Thus a protein array comprising a surface according to the invention may subsist as series of separate solid phase surfaces, such as beads carrying different proteins, the array being formed by the spatially defined pattern or arrangement of the separate surfaces in the experiment.

Preferably the surface coating is capable of resisting non-specific protein absorption. The surface coating can be porous or non-porous in nature. In addition, in a preferred embodiment the surface coating provides a specific

interaction with the marker moiety on each protein moiety either directly or indirectly (e.g. through a protein or peptide or nucleic acid bound to the surface). An embodiment of the invention described in the examples below uses SAM2™ membrane (Promega, Madison, Wisconsin, USA) as the capture surface, although a variety of other surfaces can be used, as well as surfaces in microarray or microwell formats as known in the art.

A protein moiety is a protein or a polypeptide encoded by a DNA sequence which is generally a gene or a naturally occurring variant of the gene. The protein moiety may take the form of the encoded protein, or may comprise additional amino acids (not originally encoded by the DNA sequence from which it is derived) to facilitate attachment to the array or analysis in an assay. In the case of the protein having only the amino acid sequence encoded by the naturally occurring gene, without additional sequence, such proteins may be attached to the array by way of a common feature between the variants. For example, a set of variant proteins may be attached to the array via a binding protein or an antibody which is capable of binding an invariant or common part of the individual proteins in the set. Preferably, protein moieties according to the invention are proteins tagged (via the combination of the protein encoding DNA sequence with a tag encoding DNA sequence) at either the N- or C-terminus with a marker moiety to facilitate attachment to the array.

Each position in the pattern of an array can contain, for example, either:

- a sample of a single protein type (in the form of a monomer, dimer, trimer, tetramer or higher multimer) or
- a sample of a single protein type bound to an interacting molecule (for example, nucleic acid molecule, antibody, other protein or small

molecule. The interacting molecule may itself interact with further molecules. For example, one subunit of an heteromeric protein may be attached to the array and a second subunit or complex of subunits may be tethered to the array via interaction with the attached protein subunit. In turn the second subunit or complex of subunits may then interact with a further molecule, e.g. a candidate drug or an antibody) or

- a sample of a single protein type bound to a synthetic molecule (e.g. peptide, chemical compound) or
- a sample of two different variant proteins or "haplotype proteins", for example each possessing a different complement of mutations or polymorphisms, e.g. "protein 1" is derived from a DNA sequence carrying SNP "A" and a 3 base pair deletion "X" whilst "protein 2" is derived from a DNA sequence carrying SNP "A", SNP "B" and a 3 base pair insertion "Y". Such an arrangement is capable of mimicking the heterozygous presence of two different protein variants in an individual.

Preferably the protein moiety at each position is substantially pure but in certain circumstances mixtures of between 2 and 100 different protein moieties can be present at each position in the pattern of an array of which at least one is tagged. Thus the proteins derived from the expression of more than one variant DNA sequence may be attached a single position for example, for the purposes of initial bulk screening of a set of variants to determine those sets containing variants of interest.

An embodiment of the invention described in the examples below uses a biotin tag to purify the proteins on the surface, however, the functionality of the array is independent of tag used.

"Naturally occurring variants of a DNA sequence of interest" are defined herein as being protein-encoding DNA sequences which share a common sequence and which differ only by one or more naturally occurring (i.e. present in a population and not introduced artificially) single nucleotide polymorphisms, deletions or insertions or those sequences which are considered to be haplotypes (a haplotype being a combination of variant features on a chromosome, usually within the context of a particular gene). Generally such DNA sequences are derived from the same gene in that they map to a common chromosomal locus and encode similar proteins, which may possess different phenotypes. In other words, such variants are generally naturally occurring versions of the same gene comprising one or more mutations, or their synthetic equivalents, which whilst having different codons, encode the same "wild-type" or variant proteins as those know to occur in a population.

Usefully, DNA molecules having all known mutations in a population are used to produce a set of protein moieties which are attached to the arrays of the invention. Optionally, the array may comprise a subset of variant proteins derived from DNA molecules possessing a subset of mutations, for example all known germ-line, or inheritable mutations or a subset of clinically relevant or clinically important mutations. Related DNA molecules as defined herein are related by more than just a common tag sequence introduced for the purposes or marking the resulting expressed protein. It is the sequence additional to such tags which is relevant to the relatedness of the DNA molecules. The related sequences are generally the natural coding sequence of a gene and variant forms caused by mutation. In practice the arrays of the invention carry protein moieties which are derived from DNA molecules which differ, i.e. are mutated

at 1 to 10, 1 to 7, 1 to 5, 1 to 4, 1 to 3, 1 to 2 or 1 discrete locations in the sequence of one DNA molecule relative to another, or more often relative to the wild-type coding sequence (or most common variant in a population). The difference or mutation at each discrete sequence location (for example a discrete location such as "base-pair 342" (the location can be a single base) or "base-pair 502 to base-pair 525" (the location can be a region of bases)) may be a point mutation such as a base change, for example the substitution of "A" for "G". This may lead to a "mis-sense" mutation, where one amino acid in the wild type sequence is replaced by different amino acid. A "single nucleotide polymorphism" is a mutation of a single nucleotide. Alternatively the mutation may be a deletion or insertion of 1 to 200, 1 to 100, 1 to 50, 1 to 20 or 1 to 10 bases. To give an example, insertional mutations are found in "triplet repeat" disorders such as Huntington's Disease – protein variants corresponding to such insertional mutations can be derived from various mutant forms of the gene and attached to the array to permit investigation of their phenotypes.

Thus, it is envisaged that proteins derived from related DNA molecules can be quite different in structure. For example a related DNA molecule which has undergone a mutation which truncates it, introduces a frame-shift or introduces a stop codon part-way through the wild–type coding sequence may produce a smaller or shorter protein product. Likewise mutation may cause the variant protein to have additional structure, for example a repeated domain or a number of additional amino acids either at the termini of the protein or within the sequence of the protein. Such proteins, being derived from related DNA sequences, are included within the scope of the invention.

As stated above, also included within the scope of the invention are arrays carrying protein moieties encoded by synthetic equivalents of a wild type gene (or a naturally occurring variant thereof) of a DNA sequence of interest.

Also included within the scope of the invention are arrays carrying protein moieties derived from related DNA molecules which, having variant i.e. mutated sequences, give rise to products which undergo differential pre-translational processing (e.g., alternatively spliced transcripts) or differential post-translational processing (e.g. glycosylation occurs at a particular amino acid in one expressed protein, but does not occur in another expressed protein due a codon change in the underlying DNA sequence causing the glycosylated amino acid to be absent).

Generally, related DNA molecules according to the invention are derived from genes which map to the same chromosomal locus, i.e. the related DNA molecules are different versions of the same protein coding sequence derived from a single copy of a gene, which differ as a result of natural mutation.

The wild-type (or the protein encoded by the most common variant DNA sequence in a population) of the protein is preferably included as one of the protein moieties on the array to act as a reference by which the relative activities of the proteins derived from related DNA molecules can be compared. The output of the assay indicates whether the related DNA molecule comprising a mutated gene encodes:

(1) a protein with comparable function to the wild-type protein

(2) a protein with lower or higher levels of function than the wild-type

(3) a protein with no detectable function

(4) a protein with altered post-translational modification patterns

(5) a protein with an activity that can be modified by addition of an extra component (e.g. peptide, antibody or small molecule drug candidate).

(6) a protein with an activity that can be modified by post-translational modification for example *in situ* on the chip, for example phosphorylation.

(7) a protein with an altered function under different environmental conditions in the assay, for example ionic strength, temperature or pH.

The protein moieties of the arrays of the present invention can comprise proteins associated with a disease state, drug metabolism, or may be uncharacterised. In one preferred embodiment the protein moieties encode wild type p53 and allelic variants thereof. In another preferred embodiment the arrays comprises protein moieties which encode a drug metabolising enzyme, preferably wild type p450 and allelic variants thereof.

The number of protein variants attached to the arrays of the invention will be determined by the number of variant coding sequences that occur naturally or that are of sufficient experimental, commercial or clinical interest to generate artificially. An array carrying a wild type protein and a single variant would be of use to the investigator. However in practice and in order to take advantage of the suitability of such arrays for high throughput assays, it is envisaged that 1 to 10000, 1 to 1000, 1 to 500, 1 to 400, 1 to 300, 1 to 200, 1 to 100, 1 to 75, 1 to 50, 1 to 25, 1 to 10 or 1 to 5 related DNA molecules are represented by their encoded proteins on an array. For example, in the case of the gene for p53 (the subject of one of the Examples described herein) there are currently about 50 known germ-line or inheritable mutations and more than 1000 known somatic mutations. An individual may of course inherit two different germ-line

mutations. Thus a p53 variant protein array might carry proteins derived from the 50 germ-line mutations each isolated at a different location, proteins from a clinically relevant subset of 800 somatic coding mutations (where a protein can be expressed) each isolated at a different location (or in groups of 10 at each location) and all possible pair-wise combinations of the 50 germ-line mutations each located at a different location. It can therefore be seen that an array of the invention can usefully represent individual DNA molecules containing more than 1000 different naturally occurring mutations and can accordingly carry many more, for example 10000 or more, separate discrete samples or "spots" of the protein variants derived therefrom either located alone or in combination with other variants.

In a second aspect, the invention provides a method of making a protein array comprising the steps of

a) providing DNA coding sequences which are derived from two or more naturally occurring variants of a DNA sequence of interest

b) expressing said coding sequences to provide one or more individual proteins

c) purifying said proteins

d) depositing said proteins at spatially defined locations on a surface to give an array.

Steps c) and d) are preferably combined in a single step. This can be done by means of "surface capture" by which is meant the simultaneous purification and isolation of the protein moiety on the array via the incorporated tag as described in the examples below. Furthermore, step c) may be optional as it is not necessary for the protein preparation to be pure at the location of the isolated

tagged protein – the tagged protein need not be separated from the crude lysate of the host production cell if purity is not demanded by the assay in which the array takes part.

The DNA molecules which are expressed to produce the protein moieties of the array can be generated using techniques known in the art (for example see Current Protocols in Molecular Biology, Volume 1, Chapter 8, Edited by Ausubel, FM, Brent, R, Kingston, RE, Moore, DD, Siedman, JG, Smith, JA, and Struhl, K). The ease of *in vitro* manipulation of cloned DNA enables mutations, for example SNPs, to be generated by standard molecular biological techniques such as PCR mutagenesis using the wild-type gene as a template. Therefore, only knowledge of the identity of the mutation, for example SNP (often available in electronic databases), and not the actual mutation containing DNA molecule, is required for protein array fabrication. The wild-type gene, encoding the protein of interest, is first cloned into a DNA vector for expression in a suitable host. It will be understood by those skilled in the art that the expression host need not be limited to *E. coli* – yeast, insect or mammalian cells can be used. Use of a eukaryotic host may be desirable where the protein under investigation is known to undergo post-translational modification such as glycosylation. Following confirmation of expression and protein activity, the wild-type gene is mutated to introduce the desired SNPs. The presence of the SNP is confirmed by sequencing following re-cloning.

To make the array, clones can be grown in microtiter plate format (but not exclusively) allowing parallel processing of samples in a format that is convenient for arraying onto slides or plate formats and which provides a high-throughput format. Protein expression is induced and clones are subsequently

processed for arraying. This can involve purification of the proteins by affinity chromatography, or preparation of lysates ready for arraying onto a surface which is selective for the recombinant protein ('surface capture'). Thus, the DNA molecules may be expressed as fusion proteins to give protein moieties tagged at either the N- or C- terminus with a marker moiety. As described herein, such tags may be used to purify or attach the proteins to the surface or the array. Conveniently and preferably, the protein moieties are simultaneously purified from the expression host lysate and attached to the array by means of the marker moiety. The resulting array of proteins can then be used to assay the functions of all proteins in a parallel, and therefore high-throughput manner.

In a third aspect, the invention provides a method of simultaneously determining the relative properties of members of a set of protein moieties derived from related DNA molecules, comprising the steps of: providing an array as herein described, bringing said array into contact with a test substance, and observing the interaction of the test substance with each set member on the array.

In one embodiment, the invention provides a method of screening a set of protein moieties derived from related DNA molecules for compounds (for example, a small organic molecule) which restore or disrupt function of a protein, which may reveal compounds with therapeutic advantages or disadvantages for a subset of the population carrying a particular SNP or other mutation. In other embodiments the test substance may be:

- a protein for determining relative protein:protein interactions within a set of protein moieties derived from related DNA molecules

- a nucleic acid molecule for determining relative protein:DNA or protein:RNA interactions
- a ligand for determining relative protein:ligand interactions

Results obtained from the interrogation of arrays of the invention can be quantitative (e.g. measuring binding or catalytic constants K_D & K_M), semi-quantitative (e.g. normalising amount bound against protein quantity) or qualitative (e.g. functional vs. non-functional). By quantifying the signals for replicate arrays where the ligand is added at several (for example, two or more) concentrations, both the binding affinities and the active concentrations of protein in the spot can be determined. This allows comparison of SNPs with each other and the wild-type. This level of information has not been obtained previously from arrays. Exactly the same methodology could be used to measure binding of drugs to arrayed proteins.

For example, quantitative results, K_D and B_{max}, which describe the affinity of the interaction between ligand and protein and the number of binding sites for that ligand respectively, can be derived from protein array data. Briefly, either quantified or relative amounts of ligand bound to each individual protein spot can be measured at different concentrations of ligand in the assay solution. Assuming a linear relationship between the amount of protein and bound ligand, the (relative) amount of ligand bound to each spot over a range of ligand concentrations used in the assay can be fitted to equation 1, rearrangements or derivations.

$$\text{Bound ligand} = B_{max} / ((K_D/[L])+1) \qquad \text{(Equation 1)}$$

[L] = concentration of ligand used in the assay

Preferred features of each aspect of the invention are as defined for each other aspect, *mutatis mutandis*.

Further features and details of the invention will be apparent from the following description of specific embodiments of a protein array, a p53 protein SNP array and a p450 array, and its use in accordance with the invention which is given by way of example with reference to the accompanying drawings, in which:-

Figure 1 shows p53 mutant panel expression. *E. coli* cells containing plasmids encoding human wild type p53 or the indicated mutants were induced for 4h at 30 C. Cells were lysed by the addition of lysozyme and Triton X100 and cleared lysates were analysed by Western blot. A band corresponding to full length his-tagged, biotinylated p53 runs at around 70kDa.

Figure 2 shows a gel shift assay to demonstrate DNA binding function of E.coli expressed p53. 1ul of cleared E.coli lysate containing wild type p53 (wt) or the indicated mutant was combined with 250nM DIG-labelled DNA and 0.05mg/ml polydI/dC competitor DNA. The –ve control contained only DNA. Bound and free DNA was separated through a 6% gel (NOVEX), transferred to positively charged membrane (Roche) and DIG-labelled DNA detected using an anti-DIG HRP conjugated antibody (Roche). The DNA:p53 complex is indicated by an arrow.

Figure 3 shows microarray data for the p53 DNA binding assay. Lysates were arrayed in a 4x4 pattern onto streptavidin capture membrane as detailed in A) and

probed with B) Cy3-labelled anti-histidine antibody or C) Cy3-labelled GADD45 DNA, prior to scanning in an Affymetrix 428 array scanner.

Figure 4 shows CKII phosphorylation of p53. 2ul of E.coli lysate containing p53 wild type (wt) or the indicated mutant protein were incubated with or without casein kinase II in a buffer containing ATP for 30min at 30 C. Reactions were Western blotted and phosphorylation at serine 392 detected using a phosphorylation specific antibody.

Figure 5 shows microarray data for the CKII phosphorylation assay. The p53 array was incubated with CKII and ATP for 1h at 30 C and analysed for phosphorylation at serine 392. Phosphorylation was detected for all proteins on the array except for the truncation mutants Q136X, R196X, R209X, R213X, R306X and for the amino acid mutants L344P and S392A.

Figure 6 shows a solution phase MDM2 interaction assay. 10ul of p53 containing lysate was incubated with 10ul of MDM2 containing lysate and 20ul anti-FLAG agarose in a total volume of 500ul. After incubation for 1h at room temperature the anti-FLAG agarose was collected by centrifugation, washed extensively and bound proteins analysed by Western blotting. P53 proteins were detected by Strep/HRP conjugate.

Figure 7 shows microarray data for MDM2 interaction. The p53 array was incubated with purified Cy3-labelled MDM2 protein for 1h at room temperature and bound MDM2 protein detected using a DNA array scanner (Affymetrix). MDM2 protein bound to all members of the array apart from the W23A and W23G mutants.

Figure 8a shows replicate p53 microarrays incubated in the presence of ^{33}P labelled duplex DNA, corresponding to the sequence of the GADD45 promoter element, at varying concentrations and imaged using a phosphorimager so individual spots could be quantified.

Figure 8B shows DNA binding to wild-type p53 (high affinity), R273H (low affinity) and L344P (non-binder) predicting a wild-type affinity of 7 nM.

Figure 9A shows a plasmid map of pBJW102.2 for expression of C-terminal BCCP hexa-histidine constructs.

Figure 9B shows the DNA sequence of pBJW102.2

Figure 9C shows the cloning site of pBJW102.2 from start codon. Human P450s, NADPH-cytochrome P450 reductase, and cytochrome b5 ORFs, and truncations thereof, were ligated to a *Dra*III / *Sma*I digested vector of pBJW102.2.

Figure 10A shows a vector map of pJW45

Figure 10B shows the sequence of the vector pJW45

Figure 11A shows the DNA sequence of Human P450 3A4 open reading frame.

Figure 11B. shows the amino acid sequence of full length human P450 3A4.

Figure 12A shows the DNA sequence of human P450 2C9 open reading frame.

Figure 12B shows the amino acid sequence of full length human P450 2C9

Figure 13A shows the DNA sequence of human P450 2D6 open reading frame.

Figure 13B shows the amino acid sequence of full length human P450 2D6.

Figure 14 shows a western blot and coomassie-stained gel of purification of cytochrome P450 3A4 from *E. coli*. Samples from the purification of cytochrome P450 3A4 were run on SDS-PAGE, stained for protein using coomassie or Western blotted onto nitrocellulose membrane, probed with streptavidin-HRP conjugate and visualised using DAB stain:

Lanes 1: Whole cells

Lanes 2: Lysate

Lanes 3: Lysed *E. coli* cells

Lanes 4: Supernatant from *E. coli* cell wash

Lanes 5: Pellet from *E. coli* cell wash

Lanes 6: Supernatant after membrane solublisation

Lanes 7: pellet after membrane solublisation

Lanes 8: molecular weight markers: 175, 83, 62, 48, 32, 25, 16.5, 6.5 Kda

Figure 15 shows the Coomassie stained gel of Ni-NTA column purification of cytochrome P450 3A4. Samples from all stages of column purification were run on SDS-PAGE:

Lane 1: Markers 175, 83, 62, 48, 32, 25, 16.5, 6.5 KDa

Lane 2: Supernatant from membrane solublisation

Lane 3: Column Flow-Through

Lane 4: Wash in buffer C

Lane 5: Wash in buffer D

Lanes 6&7: Washes in buffer D + 50 mM Imidazole

Lanes 8 - 12: Elution in buffer D + 200 mM Imidazole

Figure 16 shows the assay of activity for cytochrome P450 2D6 in a reconstitution assay using the substrate AMMC. Recombinant, tagged CYP2D6 was compared with a commercially available CYP2D6 in terms of ability to turnover AMMC after reconstitution in liposomes with NADPH-cytochrome P450 reductase.

Figure 17 shows the rates of resorufin formation from BzRes by cumene hydrogen peroxide activated cytochrome P450 3A4. Cytochrome P450 3A4 was assayed in solution with cumene hydrogen peroxide activation in the presence of increasing concentrations of BzRes up to 160 μM.

Figure 18 shows the equilibrium binding of $[^3H]$ketoconazole to immobilised CYP3A4 and CYP2C9. In the case of CYP3A4 the data points are the means ± standard deviation, of 4 experiments. Non-specific binding was determined in the presence of 100μM ketoconazole (data not shown).

Figure 19 shows the chemical activation of tagged, immobilised P450 involving conversion of DBF to fluorescein by CHP activated P450 3A4 immobilised on a streptavidin surface.

Figure 20 shows the stability of agarose encapsulated microsomes. Microsomes containing cytochrome P450 2D6 plus NADPH-cytochrome P450 reductase and cytochrome b5 were diluted in agarose and allowed to set in 96 well plates. AMMC turnover was measured immediately and after two and seven days at 4°C.

Figure 21 shows the turnover of BzRes by cytochrome P450 3A4 isoforms. Cytochrome P450 3A4 isoforms WT, *1, *2, *3, *4, *5 & *15, (approximately 1 μg) were incubated in the presence of BzRes (0 – 160 μM) and cumene hydrogen peroxide (200 μM) at room temperature in 200 mM KPO_4 buffer pH 7.4. Formation of resorufin was measured over time and rates were calculated from progress curves. Curves describing conventional Michaelis-Menton kinetics were fitted to the data.

Figure 22 shows the inhibition of cytochrome P450 3A4 isoforms by ketoconazole. Cytochrome P450 3A4 isoforms WT, *1, *2, *3, *4, *5 & *15, (approximately 1 μg) were incubated in the presence of BzRes (50 μM), Cumene hydrogen peroxide (200 μM) and ketoconazole (0, 0.008, 0.04, 0.2, 1, 5 μM) at room temperature in 200 mM KPO_4 buffer pH 7.4. Formation of resorufin was measured over time and rates were calculated from progress curves. IC_{50} inhibition curves were fitted to the data.

EXAMPLES

Example 1: Use of a protein array for functional analysis of proteins encoded by SNP-containing genes – the p53 protein SNP array

Mutations in the tumour suppresser protein p53 have been associated with around 50% of cancers, and more than a thousand SNPs of this gene have been observed. Mutations of the p53 gene in tumour cells (somatic mutation), or in the genome of families with a predisposition to cancer (germline mutation), provide an association between a condition and genotype, but no molecular mechanism. To demonstrate the utility of protein arrays for functional characterisation of coding SNPs, the

Inventors have arrayed wild type human p53 together with 46 germline mutations (SNPs). The biochemical activity of these proteins can then be compared rapidly and in parallel using small sample volumes of reagent or ligand. The arrayed proteins are shown to be functional for DNA binding, phosphorylated post-translationally "on-chip" by a known p53 kinase, and can interact with a known p53-interacting protein, MDM2. For many of these SNPs, this is the first functional characterisation of the effect of the mutation on p53 function, and illustrates the usefulness of protein microarrays in analysing biochemical activities in a massively parallel fashion.

Materials and Methods for construction of p53 SNP array.

Wild type p53 cDNA was amplified by PCR from a HeLa cell cDNA library using primers P53F (5' atg gag gag ccg cag tca gat cct ag 3') and P53R (5' gat cgc ggc cgc tca gtc agg ccc ttc tg 3') and ligated into an *E.coli* expression vector downstream of sequence coding for a poly Histidine-tag and the BCCP domain

from the *E.coli AccB* gene. The ligation mix was transformed into chemically competent XL1Blue cells (Stratagene) according to the manufacturer's instructions. The p53 cDNA sequence was checked by sequencing and found to correspond to wild type p53 protein sequence as contained in the SWISS-PROT entry for p53 [Accession No. P04637].

Construction of p53 mutant panel

Mutants of p53 were made by using the plasmid containing the wild type p53 sequence as template in an inverse PCR reaction. Primers were designed such that the forward primer was 5' phosphorylated and started with the single nucleotide polymorphism (SNP) at the 5' end, followed by 20-24 nucleotides of p53 sequence. The reverse primer was designed to be complementary to the 20-24 nucleotides before the SNP. PCR was performed using Pwo polymerase which generated blunt ended products corresponding to the entire p53-containing vector. PCR products were gel purified, ligated to form circular plasmids and parental template DNA was digested with restriction endonuclease DpnI (New England Biolabs) to increase cloning efficiency. Ligated products were transformed into XL1Blue cells, and mutant p53 genes were verified by sequencing for the presence of the desired mutation and the absence of any secondary mutation introduced by PCR.

Expression of p53 in E.coli

Colonies of XLIBlue cells containing p53 plasmids were inoculated into 2 ml of LB medium containing ampicillin (70 micrograms /ml) in 48 well blocks (QIAGEN) and grown overnight at 37 °C in a shaking incubator. 40 µl of overnight culture was used to inoculate another 2 ml of LB/ampicillin in 48 well blocks and grown at 37 °C until an optical density (600nm) of ~0.4 was

reached. IPTG was then added to 50 μM and induction continued at 30 °C for 4 hours. Cells were then harvested by centrifugation and cell pellets stored at –80 °C. For preparation of protein, cell pellets were thawed at room temperature and 40 μl of p53 buffer (25 mM HEPES pH 7.6, 50 mM KCl, 10% glycerol, 1 mM DTT, 1 mg/ml bovine serum albumin, 0.1% Triton X100) and 10 μl of 4 mg/ml lysozyme were added and vortexed to resuspend the cell pellet. Lysis was aided by incubation on a rocker at room temperature for 30 min before cell debris was collected by centrifugation at 13000 rpm for 10 min at 4 °C. The cleared supernatant of soluble protein was removed and used immediately or stored at –20 °C.

Western blotting

Soluble protein samples were boiled in SDS containing buffer for 5 min prior to loading on 4-20% Tris-Glycine gels (NOVEX) and run at 200 V for 45 min. Protein was transferred onto PVDF membrane (Hybond-P, Amersham) and probed for the presence of various epitopes using standard techniques. For detection of the histidine-tag, membranes were blocked in 5% Marvel /PBST and anti-RGSHis antibody (QIAGEN) was used as the primary antibody at 1/1000 dilution. For detection of the biotin tag, membranes were blocked in Superblock /TBS (Pierce) and probed with Streptavidin-HRP conjugate (Amersham) at 1/2000 dilution in Superblock/TBS/0.1% Tween20. The secondary antibody for the RGSHis antibody was anti-mouse IgG (Fc specific) HRP conjugate (Sigma) used at 1/2000 dilution in Marvel/PBST. After extensive washing, bound HRP conjugates were detected using either ECLPlus (Amersham) and Hyperfilm ECL (Amersham) or by DAB staining (Pierce).

DNA gel shift assay

DNA binding function of expressed p53 was assayed using a conventional gel shift assay. Oligos DIGGADD45A (5'DIG-gta cag aac atg tct aag cat gct ggg gac-3') and GADD45B (gtc ccc agc atg ctt aga cat gtt ctg tac 3') were annealed together to give a final concentration of 25 μM dsDNA. Binding reactions were assembled containing 1 μl of cleared lysate, 0.2 μl of annealed DIG-labelled GADD45 oligos and 1 μl of polydI/dC competitor DNA (Sigma) in 20 μl of p53 buffer. Reactions were incubated at room temperature for 30 min, chilled on ice and 5 μl loaded onto a pre-run 6% polyacrylamide/TBE gel (NOVEX). Gels were run at 100 V at 4 °C for 90 min before being transferred onto positively charged nitrocellulose (Roche). Membranes were blocked in 0.4% Blocking Reagent (Roche) in Buffer I (100 mM maleic acid, 150 mM NaCl, pH 7.0) for 30 min and probed for presence of DIG-labelled DNA with anti-DIG Fab fragments conjugated to HRP (Roche). Bound HRP conjugates were detected using ECLPlus and Hyperfilm ECL (Amersham).

p53 phosphorylation assay

Phosphorylation of p53 was performed using purified casein kinase II (CKII, Sigma). This kinase has previously been shown to phosphorylate wild type p53 at serine 392. Phosphorylation reactions contained 2 μl of p53 lysate, 10 mM $MgCl_2$, 100 μM ATP and 0.1U of CKII in 20 μl of p53 buffer. Reactions were incubated at 30 °C for 30 min, reaction products separated through 4-20% NOVEX gels and transferred onto PVDF membrane. Phosphorylation of p53 was detected using an antibody specific for phosphorylation of p53 at serine 392 (Cell Signalling Technology), used at 1/1000 dilution in Marvel/TBST. Secondary antibody was an anti-rabbit HRP conjugate (Cell Signalling Technology), used at 1/2000 dilution.

MDM2 interaction assay

The cDNA for the N-terminal portion of MDM2 (amino acids 17-127) was amplified from a cDNA library and cloned downstream of sequences coding for a His-tag and a FLAG-tag in an *E. coli* expression vector. Plasmids were checked by sequencing for correct MDM2 sequence and induction of *E. coli* cultures showed expression of a His and FLAG tagged soluble protein of the expected size. To test for interaction between MDM2 and the p53 mutant panel, binding reactions were assembled containing 10μl p53 containing lysate, 10μl MDM2 containing lysate, 20μl anti-FLAG agarose in 500μl phosphate buffered saline containing 300mM NaCl, 0.1% Tween20 and 1% (w/v) bovine serum albumin. Reactions were incubated on a rocker at room temperature for 1 hour and FLAG bound complexes harvested by centrifugation at 5000rpm for 2min. After extensive washing in PBST, FLAG bound complexes were denatured in SDS sample buffer and Western blotted. Presence of biotinylated p53 was detected by Streptavidin/HRP conjugate.

p53 microarray fabrication and assays

Cleared lysates of the p53 mutant panel were loaded onto a 384 well plate and printed onto SAM2™ membrane (Promega, Madison, Wisconsin, USA) using a custom built robot (K-Biosystems, UK) with a 16 pin microarraying head. Each lysate was spotted 4 times onto each array, and each spot was printed onto 3 times. After printing, arrays were wet in p53 buffer and blocked in 5% Marvel/p53 buffer for 30min. After washing 3 x 5min in p53 buffer, arrays were ready for assay.

For DNA binding assay, 5μl of annealed Cy3-labelled GADD45 oligo was added to 500μl p53 buffer. The probe solution was washed over the array at

room temperature for 30min, and washed for 3 x 5min in p53 buffer. Arrays were then dried and mounted onto glass slides for scanning in an Affymetrix 428 array scanner. Quantification of Cy3 scanned images was accomplished using ImaGene software.

For the phosphorylation assay, 10μl CKII was incubated with the arrays in 320μl p53 buffer and 80μl Mg/ATP mix at 30°C for 30min. Arrays were then washed for 3 x 5min in TBST and anti-phosphoserine 392 antibody added at 1/1000 dilution in Marvel/TBST for 1h. After washing for 3 x 5min in TBST, anti-rabbit secondary antibody was added at 1/2000 dilution for 1h. Bound antibody was detected by ECLPlus and Hyperfilm.

For the MDM2 interaction assay, 1μl of purified Cy3 labelled MDM2 protein was incubated with the arrays in 500μl PBS/300mM NaCl/0.1% Tween20/1% BSA for 1h at room temperature. After washing for 3 x 5min in the same buffer, arrays were dried, mounted onto glass slides and analysed for Cy3 fluorescence as for the DNA binding assay.

Results

Expression of p53 in E.coli and construction of mutant panel

The full length p53 open reading frame was amplified from a Hela cell cDNA library by PCR and cloned downstream of the tac promoter in vector pQE80L into which the BCCP domain from the E.coli gene ACCB had already been cloned. The resultant p53 would then be His and biotin tagged at its N-terminus, and figure 1 shows Western blot analysis of soluble protein from induced E.coli cultures. There is a clear signal for His-tagged, biotinylated protein at around 66kDa, and a band of the same size is detected by the p53 specific antibody pAb1801 (data not shown). The plasmid encoding this protein was fully sequenced and shown to be wild type p53 cDNA sequence. This plasmid was

used as the template to construct the mutant panel, and figure 1 also shows analysis of the expression of a selection of those mutants, showing full length protein as expected for the single nucleotide polymorphisms, and truncated proteins where the mutation codes for a STOP codon. The mutants were also sequenced to confirm presence of the desired mutation and absence of any secondary mutations.

Although the Inventors have used His and biotin tags in this example of a SNP array, other affinity tags (eg FLAG, myc, VSV) can be used to enable purification of the cloned proteins. Also an expression host other than E. coli can be used (eg. yeast, insect cells, mammalian cells) if required.
Also, although this array was focussed on the naturally occurring germline SNPs of p53, other embodiments are not necessarily restricted to naturally occurring SNPs ("synthetic" mutants) or versions of the wild type protein which contain more than one SNP. Other embodiments can contain versions of the protein which are deleted from either or both ends (a nested-set). Such arrays would be useful in mapping protein:ligand interactions and delineating functional domains of unknown proteins.

E. coli expressed p53 is functional for DNA binding
To demonstrate functionality of our p53, the Inventors performed electrophoretic mobility shift assays using a DNA oligo previously shown to be bound by p53. Figure 2 shows an example result from these gel shift assays, showing DNA binding by wild type p53 as well as mutants R72P, P82L and R181C. The first 2 mutants would still be expected to bind DNA as these mutations are outside of the DNA binding domain of p53. Having demonstrated DNA binding using a conventional gel based assay, the Inventors then wanted

to show the same function for p53 arrayed on a surface. Figure 3C shows the result of binding Cy3-labelled DNA to the p53 mutant panel arrayed onto SAM2™ membrane (Promega, Madison, Wisconsin, USA). Although the Inventors have used SAM2™ membrane in this example of a SNP array, other surfaces which can be used for arraying proteins onto include but are not restricted to glass, polypropylene, polystyrene, gold or silica slides, polypropylene or polystyrene multi-well plates, or other porous surfaces such as nitrocellulose, PVDF and nylon membranes. The SAM2™ membrane specifically captures biotinylated molecules and so purifies the biotinylated p53 proteins from the mutant panel cell lysates. After washing unbound DNA from the array, bound DNA was visualised using an Affymetrix DNA array scanner. As can be seen from figure 3, the same mutants which bound DNA in the gel shift assay also bound the most DNA when arrayed on a surface. Indeed, for a DNA binding assay the microarray assay appeared to be more sensitive than the conventional gel shift assay. This is probably because in a gel shift assay the DNA:protein complex has to remain bound during gel electrophoresis, and weak complexes may dissociate during this step. Also the 3-dimensional matrix of the SAM2™ membrane used may have a caging effect. The amount of p53 protein is equivalent on each spot, as shown by an identical microarray probed for His-tagged protein (figure 3B).

Use of the p53 array for phosphorylation studies

To exemplify the study of the effect of SNPs on post-translational modifications, the Inventors chose to look at phosphorylation of the p53 array by casein kinase II. This enzyme has previously been shown to phosphorylate p53 at serine 392, and the Inventors made use of a commercially available anti-p53 phosphoserine 392 specific antibody to study this event. Figure 4 shows

Western blot analysis of kinase reactions on soluble protein preparations from p53 wild type and S392A clones. Lane 1 shows phosphorylation of wild type p53 by CKII, with a background signal when CKII is omitted from the reaction (lane 2). Lanes 3 and 4 show the corresponding results for S392A, which as expected only shows background signal for phosphorylation by CKII. This assay was then applied in a microarray format, which as can be seen from figure 5 shows phosphorylation for all of the mutant panel except the S392A mutant and those mutants which are truncated before residue 392.

Use of the p53 array to study a protein:protein interaction

To exemplify the study of a protein:protein interaction on a SNP protein array, the interaction of MDM2 with the p53 protein array was investigated. Figure 6 shows that FLAG-tagged MDM2 pulls down wild type p53 when bound to anti-FLAG agarose. However the W23A mutant is not pulled down by FLAG agarose bound MDM2, which would be expected as this residue has previously been shown to be critical for the p53/MDM2 interaction (Bottger, A., Bottger, V., Garcia-Echeverria, C., et al, J. Mol. Biol. (1997) 269: 744-756). This assay was then carried out in a microarray format, and figure 7 shows the result of this assay, with Cy3-labelled protein being detected at all spots apart from the W23A and W23G mutant spots.

The Inventors have used a novel protein chip technology to characterise the effect of 46 germline mutations on human p53 protein function. The arrayed proteins can be detected by both a His-tagged antibody and also a p53 specific antibody. This array can be used to screen for mutation specific antibodies which could have implications for p53 status diagnosis.

The Inventors were able to demonstrate functionality of the wild type protein by conventional gel based assays, and have achieved similar results performing the assays in a microarray format. Indeed, for a DNA binding assay the microarray assay appeared to be more sensitive than the conventional gel shift assay. These arrays can be stored at -20 C in 50% glycerol and have been shown to still be functional for DNA binding after 1 month (data not shown).

The CKII phosphorylation assay results are as expected, with phosphorylation being detected for all proteins which contained the serine at residue 392. This analysis can obviously be extended to a screen for kinases that phosphorylate p53, or for instance for kinases that differentially phosphorylate some mutants and not others, which could themselves represent potential targets in cancer.

The MDM2 interaction assay again shows the validity of the protein array format, with results for wild type and the p53 mutants mirroring those obtained using a more conventional pull down assay. These results also show that our protein arrays can be used to detect protein:protein interactions. Potentially these arrays can be used to obtain quantitative binding data (ie K_D values) for protein:protein interactions in a high-throughput manner not possible using current methodology. The fact that the MDM2 protein was pulled out of a crude E. coli lysate onto the array bodes well for envisioned protein profiling experiments, where for instance cell extracts are prepared from different patients, labelled with different fluorophores and both hybridised to the same array to look for differences in amounts of protein interacting species.

Indeed, in Example 2 below the applicant has gone on to demonstrate that these arrays can be used to obtain quantative data.

Example 2 Quantitative DNA binding on the p53 protein microarray

Methods

DNA-binding assays. Oligonucleotides with the GADD45 promoter element sequence (5'-gta cag aac atg tct aag cat gct ggg gac-3' and 5'-gtc ccc agc atg ctt aga cat gtt ctg tac-3') were radiolabelled with gamma ^{33}P-ATP (Amersham Biosciences, Buckinghamshire, UK) and T4 kinase (Invitrogen, Carlsbad, CA), annealed in p53 buffer and then purified using a Nucleotide Extraction column (Qiagen, Valencia, CA). The duplex oligos were quantified by UV spectrophotometry and a 2.5 fold dilution series made in p53 buffer. 500 µl of each dilution were incubated with microarrays at room temperature for 30 min, then washed three times for 5 min in p53 buffer to remove unbound DNA. Microarrays were then exposed to a phosphorimager plate (Fuji, Japan) overnight prior to scanning. ImaGene software (BioDiscovery, Marina del Rey, CA) was used to quantify the scanned images. Replicate values for all mutants at each DNA concentration were fitted to simple hyperbolic concentration-response curves $R=B_{max}/((K_d /L)+1)$, where R is the response in relative counts and L is the DNA concentration in nM.

Results

Binding of p53 to GADD45 promoter element DNA. Replicate p53 microarrays were incubated in the presence of ^{33}P labelled duplex DNA, corresponding to the sequence of the GADD45 promoter element, at varying concentrations (Fig. 8A). The microarrays were imaged using a phosphorimager and individual spots quantified. The data were normalised against a calibration curve to compensate for the non-linearity of this method of detection and

backgrounds were subtracted. Replicate values for all mutants were plotted and analysed by non-linear regression analysis allowing calculation of both K_d and B_{max} values (Table 1).

Table 1

Mutation	DNA binding				MDM2	CKII
	B_{max} (% wild-type)		K_d (nM)			
Wild-type	100	(90-110)	7	(5-10)	+	+
W23A	131	(119-144)	7	(5-10)	-	+
W23G	84	(74-94)	5	(3-9)	-	+
R72P	121	(110-132)	9	(7-13)	+	+
P82L	70	(63-77)	7	(5-10)	+	+
M133T	ND				+	+
Q136X	No binding				+	-
C141Y	ND				+	+
P151S	ND				+	+
P152L	31	(23-38)	18	(9-37)	+	+
G154V	ND				+	+
R175H	ND				+	+
E180K	31	(21-41)	12	(4-35)	+	+
R181C	88	(81-95)	11	(8-13)	+	+
R181H	48	(40-57)	11	(6-21)	+	+
H193R	21	(16-26)	22	(11-42)	+	+
R196X	No binding				+	-
R209X	No binding				+	-
R213X	No binding				+	-
P219S	21	(14-30)	10	(3-33)	+	+
Y220C	ND				+	+
S227T	101	(94-110)	7	(5-9)	+	+
H233N	60	(52-68)	5	(3-8)	+	+
H233D	70	(58-84)	7	(3-14)	+	+
N235D	32	(25-40)	27	(15-49)	+	+
N235S	46	(36-56)	9	(4-20)	+	+
S241F	38	(30-47)	19	(10-37)	+	+
G245C	ND				+	+
G245S	44	(38-51)	11	(7-18)	+	+
G245D	ND				+	+
R248W	107	(95-120)	12	(8-17)	+	+
R248Q	85	(77-95)	17	(12-23)	+	+
I251M	ND				+	+
L252P	22	(12-32)	16	(4-63)	+	+
T256I	32	(22-41)	14	(6-34)	+	+
L257Q	26	(19-35)	17	(7-44)	+	+
E258K	ND				+	+
L265P	ND				+	+
V272L	ND				+	+
R273C	70	(56-85)	20	(11-37)	+	+
R273H	59	(40-79)	54	(27-106)	+	+
P278L	ND				+	+
R280K	54	(40-70)	21	(9-46)	+	+
E286A	32	(23-41)	22	(10-46)	+	+
R306X	No binding				+	-
R306P	90	(81-100)	7	(5-11)	+	+
G325V	73	(67-79)	7	(5-10)	+	+
R337C	88	(80-95)	6	(4-8)	+	+
L344P	No binding				+	-
S392A	121	(107-136)	10	(6-14)	+	-

Figure 8B shows DNA binding to wild-type p53 (high affinity), R273H (low affinity) and L344P (non-binder) predicting a wild-type affinity of 7 nM.

Discussion

DNA binding. Quantitative analysis of the DNA binding data obtained from the microarrays yielded both affinities (K_d) and relative maximum binding values (B_{max}) for wild-type and mutant p53. Protein function microarrays have not previously been used in this way and this data therefore demonstrate their usefulness in obtaining this quality and amount of data in a parallel fashion. The approach of normalising binding data for the amount of affinity-tagged protein in the spot provides a rapid means of analysing large data sets [Zhu, H. et al. Global analysis of protein activities using proteome chips. *Science* **293**, 2101-2105 (2001).], however it takes into account neither the varying specific activity of the microarrayed protein nor whether the signal is recorded under saturating or sub-saturating conditions. The quantitative analysis carried out here allowed the functional classification of mutants into groups according to GADD45 DNA binding: those showing near wild-type affinity; those exhibiting reduced stability (low B_{max}); those showing reduced affinity (higher K_d); and those showing complete loss of activity (Table 1).

Proteins with near wild-type affinity for DNA generally had mutations located outside of the DNA-binding domain and include R72P, P82L, R306P and G325V. R337C is known to affect the oligomerisation state of p53 but at the assay temperature used here it is thought to be largely tetrameric [Davison, T.S., Yin, P., Nie, E., Kay, C. & Arrowsmith, C.H. Characterisation of the oligomerisation defects of two p53 mutants found in families with Li-Fraumeni and Li-Fraumeni like syndrome. *Oncogene* **17**, 651-656 (1998).], consistent with the affinity measured here. By contrast, total loss of binding was observed for mutations introducing premature stop codons (Q136X, R196X, R209X and

R213X) and mutations that monomerise the protein (L344P [Lomax, M.E., Barnes, D.M., Hupp, T.R., Picksley, S.M. & Camplejohn, R.S. Characterisation of p53 oligomerisation domain mutations isolated from Li-Fraumeni and Li-Fraumeni like family members. *Oncogene* **17**, 643-649 (1998).] and the tetramerisation domain deficient R306X) as expected.

Within the DNA-binding domain, the applicant found that mutations generally reduced or abolished DNA binding with the notable exceptions of R181C/H, S227T and H233N/D; these are all solvent exposed positions, distant from the protein-DNA interface and exhibit wild-type binding. Mutations R248Q/W, R273C/H and R280K, present at the protein-DNA interface, exhibited low affinities with K_d values 2-7 times higher than wild-type (Table 1) consistent with either loss of specific protein-DNA interactions or steric hindrance through sub-optimal packing of the mutated residue.

Many of the remaining mutants fall into a group displaying considerably reduced specific activities, apparent from very low B_{max} values, even when normalised according to the amount of protein present in the relevant spot. For some mutants, DNA binding was compromised to such a level that although binding was observed, it was not accurately quantifiable due to low signal to background ratios e.g. P151S and G245C. For others such as L252P, low signal intensities yielded measurable K_d values, but with wide confidence limits.

To further demonstrate the applicability of the invention to protein arrays comprising at least two protein moieties derived from naturally occurring variants of a DNA sequence of interest such as, for example, those encoding proteins from phase 1 or phase 2 drug metabolising enzymes (DME's) the invention is further exemplified with reference to a p450 array. Phase 1 DME's include the Cytochrome p450's and the Flavin mono oxygenases (FMO's) and the Phase 2 DME's, UDP–glycosyltransferase (UGTs), glutathione S

transferases (GSTs), sulfotransferases (SULTs), N -acetyltransferases (NATs), drug binding nuclear receptors and drug transporter proteins.

Preferably, the full complement, or a significant proportion of human DMEs are present on the arrays of the invention. Such an array can include (numbers in parenthesis currently described in the Swiss Prot database): all the human P450s (119), FMOs (5), UDP–glycosyltransferase (UGTs) (18), GSTs (20), sulfotransferases (SULTs) (6), N-acetyltransferases (NATs) (2), drug binding nuclear receptors (33) and drug transporter proteins (6). This protein list does not include those yet to be characterised from the human genome sequencing project, splice variants known to occur for the P450s that can switch substrate specificity or polymorphisms known to affect the function and substrate specificity of both the P450s and the phase 2 DMEs.

For example it is known that there are large differences in the frequency of occurrence of various alleles in P450s 2C9, 2D6 and 3A4 between different ethnic groups (see Tables 2, 3 and 4). These alleles have the potential to affect enzyme kinetics, substrate specificity, regio-selectivity and, where multiple products are produced, product profiles. Arrays of proteins described in this disclosure allow a more detailed examination of these differences for a particular drug and will be useful in predicting potential problems and also in effectively planning the population used for clinical trials.

Table 2. P450 2D6 Allele Frequency

P450	Allele	Mutation	Allele Frequency	Ethnic Group	Study Group	Reference
2D6	*1	W.T.	26.9% 36.4% 36% 33%	Chinese German Caucasian European	113 589 195 1344	(1) (2) (3) (4)
2D6	*2	R296C; S486T	13.4% 32.4% 29% 27.1%	Chinese German Caucasian European	113 589 195 1344	(1) (2) (3) (4)
2D6	*3	Frameshift	2% 1% 1.9%	German Caucasian European	589 195 1344	(2) (3) (4)
2D6	*4	Splicing defect	20.7% 20% 16.6% 1.2%	German Caucasian European Ethiopian	589 195 1344 115	(2) (3) (4) (5)
2D6	*5	Deletion	4% 6.9%	Caucasian European	195 1344	(3) (4)
2D6	*6	Splicing defect	0.93% 1.3%	German Caucasian	589 195	(2) (3)
2D6	*7	H324P	0.08% 0.3% 0.1%	German Caucasian European	589 195 1344	(2) (3) (4)
2D6	*9	K281del	2% 2.7%	Caucasian European	195 1344	(3) (4)
2D6	*10	P34S; S486T	50.7% 1.53% 2%	Chinese German Caucasian	113 589 195	(1) (2) (3)

			1.5% 8.6%	European Ethiopian	1344 115	(4) (5)
2D6	*12	G42R; R296C; S486T	0% 0.1%	German European	589 1344	(2) (4)
2D6	*14	P34S; G169R; R296C; S486T	0.1%	European	1344	(4)
2D6	*17	T107I; R296C; S486T	0% 0.1% 9% 34%	Caucasian European Ethiopian African	195 1344 115 388	(3) (4) (5) (6)

All other P450 allelic variants occur at a frequency of 0.1 % or less (4).

Table 3 P450 2C9 Allele Frequency

P450	Allele	Mutation	Allele Frequency	Ethnic Group	Study Group	Reference
2C9	*1	W.T.	62%	Caucasian	52	(7)
2C9	*2	R144C	17%	Caucasian	52	(7)
2C9	*3	I359L	19%	Caucasian	52	(7)
2C9	*4	I359T	x%	Japanese	X	(8)
2C9	*5	D360E	0% 3%	Caucasians African-Americans	140 120	(9) (9)
2C9	*7	Y358C	x%		X	Swiss Prot

Table 4. P450 3A4 Allele Frequency

P450	Allele	Mutation	Allele Frequency	Ethnic Group	Study Group	Reference
3A4	*1	W.T.	>80%		X	
3A4	*2	S222P	2.7% 0% 0%	Caucasian African Chinese	X x x	(10) (10) (10)
3A4	*3	M445T	1% 0.47% 4%	Chinese European Caucasian	X 213 72	(10) (11) (12)
3A4	*4	I118V	2.9%	Chinese	102	(13)
3A4	*5	P218R	2%	Chinese	102	(13)
3A4	*7	G56D	1.4%	European	213	(11)
3A4	*8	R130Q	0.33%	European	213	(11)
3A4	*9	V170I	0.24%	European	213	(11)
3A4	*10	D174H	0.24%	European	213	(11)
3A4	*11	T363M	0.34%	European	213	(11)
3A4	*12	L373F	0.34%	European	213	(11)
3A4	*13	P416L	0.34%	European	213	(11)
3A4	*15	R162Q	4%	African	72	(12)
3A4	*17	F189S	2%	Caucasian	72	(12)
3A4	*18	L293P	2%	Asian	72	(12)
3A4	*19	P467S	2%	Asian	72	(12)

References

1. Johansson, I., Oscarson, M., Yue, Q. Y., Bertilsson, L., Sjoqvist, F. & Ingelman-Sundberg, M. (1994) *Mol Pharmacol* 46, 452-9.
2. Sachse, C., Brockmoller, J., Bauer, S. & Roots, I. (1997) *Am J Hum Genet* 60, 284-95.

3. Griese, E. U., Zanger, U. M., Brudermanns, U., Gaedigk, A., Mikus, G., Morike, K., Stuven, T. & Eichelbaum, M. (1998) *Pharmacogenetics* 8, 15-26.

4. Marez, D., Legrand, M., Sabbagh, N., Guidice, J. M., Spire, C., Lafitte, J. J., Meyer, U. A. & Broly, F. (1997) *Pharmacogenetics* 7, 193-202.

5. Aklillu, E., Persson, I., Bertilsson, L., Johansson, I., Rodrigues, F. & Ingelman-Sundberg, M. (1996) *J Pharmacol Exp Ther* 278, 441-6.

6. Dandara, C., Masimirembwa, C. M., Magimba, A., Sayi, J., Kaaya, S., Sommers, D. K., Snyman, J. R. & Hasler, J. A. (2001) *Eur J Clin Pharmacol* 57, 11-7.

7. Aithal, G. P., Day, C. P., Kesteven, P. J. & Daly, A. K. (1999) *Lancet* 353, 717-9.

8. Imai, J., Ieiri, I., Mamiya, K., Miyahara, S., Furuumi, H., Nanba, E., Yamane, M., Fukumaki, Y., Ninomiya, H., Tashiro, N., Otsubo, K. & Higuchi, S. (2000) *Pharmacogenetics* 10, 85-9.

9. Dickmann, L. J., Rettie, A. E., Kneller, M. B., Kim, R. B., Wood, A. J., Stein, C. M., Wilkinson, G. R. & Schwarz, U. I. (2001) *Mol Pharmacol* 60, 382-7.

10. Sata, F., Sapone, A., Elizondo, G., Stocker, P., Miller, V. P., Zheng, W., Raunio, H., Crespi, C. L. & Gonzalez, F. J. (2000) *Clin Pharmacol Ther* 67, 48-56.

11. Eiselt, R., Domanski, T. L., Zibat, A., Mueller, R., Presecan-Siedel, E., Hustert, E., Zanger, U. M., Brockmoller, J., Klenk, H. P., Meyer, U. A., Khan, K. K., He, Y. A., Halpert, J. R. & Wojnowski, L. (2001) *Pharmacogenetics* 11, 447-58.

12. Dai, D., Tang, J., Rose, R., Hodgson, E., Bienstock, R. J., Mohrenweiser, H. W. & Goldstein, J. A. (2001) *J Pharmacol Exp Ther* 299, 825-31.

13. Hsieh, K. P., Lin, Y. Y., Cheng, C. L., Lai, M. L., Lin, M. S., Siest, J. P. & Huang, J. D. (2001) *Drug Metab Dispos* 29, 268-73.

Example 3: Cloning of wild-type *H. sapiens* cytochrome P450 enzymes CYP2C9, CYP2D6 and CYP3A4

The human cytochrome p450s have a conserved region at the N-terminus, this includes a hydrophobic region which faciliates lipid association, an acidic or 'stop transfer' region, which stops the protein being fed further into the membrane, and a partially conserved proline repeat. Three versions of the p450s were produced with deletions up to these domains, the N-terminal deletions are shown below.

Construct	Version	N-terminal Deletion
T009-C2 3A4	Proline	-34 AA
T009-C1 3A4	Stop Transfer	-25 AA
T009-C3 3A4	Hydrophobic peptide	-13 AA
T015-C2 2C9	Proline	-28 AA
T015-C1 2C9	Stop Transfer	-20 AA
T015-C3 2C9	Hydrophobic peptide	-0AA
T017-C1 2D6	Proline	-29 AA
T017-C2 2D6	Stop Transfer	-18 AA
T017-C3 2D6	Hydrophobic peptide	-0 AA

The human CYP2D6 was amplified by PCR from a pool of brain, heart and liver cDNA libraries (Clontech) using specific forward and reverse primers (T017F and T017R). The PCR products were cloned into the pMD004 expression vector, in frame with the N-terminal His-BCCP tag and using the Not1 restriction site present in the reverse primer. To convert the CYP2D6 for expression in the C-terminal tag vector pBJW102.2 (Fig. 9A&B), primers were used which incorporated an Sfi1 cloning site at the 5' end and removed the stop

codon at the 3' to allow in frame fusion with the C-terminal tag. The primers T017CR together with either T017CF1, T017CF2, or T017CF3 allowed the deletion of 29, 18 and 0 amino acids from the N-terminus of CYP2D6 respectively.

Primer sequences are as follows:

T017F:	5'-GCTGCACGCTACCCACCAGGCCCCCTG-3'.
T017R:	5'-TTGCGGCCGCTCTTCTACTAGCGGGGCACAGCACAAAGCTCATAG-3'
T017CF1:	5'-TATTCTCACTGGCCATTACGGCCGCTGCACGCTACCCACCAGGCCCCCTG-3'
T017CF2:	5'-TATTCTCACTGGCCATTACGGCCGTGGACCTGATGCACCGGCGCCAACGCTGGGCTGCACGCTACCCACCAGGCCCCCTG-3'
T017CF3:	5'-TATTCTCACTGGCCATTACGGCCATGGCTCTAGAAGCACTGGTGCCCCTGGCCGTGATAGTGGCCATCTTCCTGCTCCTGGTGGACCTGATGCACCGGCGCCAACGC-3'
T017CR:	5'-GCGGGGCACAGCACAAAGCTCATAGGG-3'

PCR was performed in a 50μl volume containing 0.5μM of each primer, 125-250μM dNTPs, 5ng of template DNA, 1x reaction buffer, 1-5 units of polymerase (Pfu, Pwo, or 'Expand long template' polymerase mix), PCR cycle = 95°C 5minutes, 95°C 30 seconds, 50-70°C 30 seconds, 72°C 4 minutes X 35 cycles, 72°C 10 minutes, or in the case of Expand 68°C was used for the extension step. PCR products were resolved by agarose gel electrophoresis, those products of the correct size were excised from the gel and subsequently purified using a gel extraction kit. Purified PCR products were then digested with either Sfi1 or Not1 and ligated into the prepared vector backbone (Fig. 9C). Correct recombinant clones were determined by PCR screening of bacterial cultures, Western blotting and by DNA sequence analysis.

CYP3A4 and CYP2C9 were cloned from cDNA libraries by a methodology similar to that of CYP2D6. Primer sequences to amplify CYP3A4 and CYP2C9 for cloning into the N-terminal vectors are as follows;

2C9

T015F: 5'-CTCCCTCCTGGCCCCACTCCTCTCCCAA-3'

T015R: 5'-TTTGCGGCCGCTCTTCTATCAGACAGGAATGAAGCACAGCCTGGTA-3'

3A4

T009F: 5'-CTTGGAATTCCAGGGCCCACACCTCTG-3'

T009R: 5'-TTTGCGGCCGCTCTTCTATCAGGCTCCACTTACGGTGCCATCCCTTGA-3'

Primers to convert the N-terminal clones for expression in the C-terminal tagging vector are as follows:

<u>3A4</u>

T009CF1: 5'-TATTCTCACTGGCCATTACGGCCTATGGAACCCATTCACATGGACTTTTTA AGAAGCTTGGAATTCCAGGGCCCACACCTCTG-3'

T009CF2: 5'-TATTCTCACTGGCCATTACGGCCCTTGGAATTCCAGGGCCCACACCTCTG-3'

T009CF3: 5'-TTCTCACTGGCCATTACGGCCCCTCCTGGCTGTCAGCCTGGTGCTCCTCTATCT ATATGGAACCCATTCACATGGACTTTTTAGG-3'

T009CR: 5'-GGCTCCACTTACGGTGCCATCCCTTGAC-3'

2C9

T015CF1: 5'-TATTCTCACTGGCCATTACGGCCAGACAGAGCTCTGGGAGAGGAAAACTCCCTC CTGGCCCCACTCCTCTCCCAG-3'

T015CF2: 5'-TATTCTCACTGGCCATTACGGCCCTCCCTCCTGGCCCCACTCCTCTCCCAG-3'

T015CR: 5'-GACAGGAATGAAGCACAGCTGGTAGAAGG-3'

The full length or Hydrophobic peptide (C3) version of 2C9 was produced by inverse PCR using the 2C9-stop transfer clone (C1) as the template and the following primers:

2C9-hydrophobic-peptide-F:

5'-CTCTCATGTTTGCTTCTCCTTTCACTCTGGAGACAGCGCTCTGGGAGAGGAAAACTC-3'

2C9-hydrophobic-peptide-R:

5'-ACAGAGCACAAGGACCACAAGAGAATCGGCCGTAAGTGCCATAGTTAATTTCTC-3'

Example 4: Cloning of NADPH-cytochrome P450 reductase

NADPH-cytochrome P450 reductase was amplified from fetal liver cDNA (Clontech), the PCR primers [NADPH reductase F1 5'-GGATCGACATATGGGAGACTCCCACGTGGACAC-3'; NADPH reductase R1 5'-CCGATAAGCTTATCAGCTCCACACGTCCAGGGAG-3'] incorporated a Nde I site at 5' and a Hind III site at the 3' of the gene to allow cloning. The PCR product was cloned into the pJW45 expression vector (Fig. 10A&B)), two stop codons were included on the reverse primer to ensure that the His-tag was not translated. Correct recombinant clones were determined by PCR screening of bacterial cultures, and by sequencing.

Example 5: Cloning of polymorphic variants of *H. sapiens* cytochrome P450s CYP2C9, CYP2D6 and CYP3A4

Once the correct wild-type CYP450s (Figs. 11, 12, & 13) were cloned and verified by sequence analysis the naturally occurring polymorphisms of 2C9, 2D6 and 3A4 shown in Table 5 were created by an inverse PCR approach (except for CYP2D6*10 which was amplified and cloned as a linear PCR product in the same way as the initial cloning of CYP2D6 described in Example 3). In each case, the forward inverse PCR primer contained a 1bp mismatch at the 5' position to substitute the wild type nucleotide for the polymorphic nucleotide as observed in the different ethnic populations.

Cytochrome P450 polymorphism	**Encoded amino acid subsitutions**
CYP2C9*1	wild-type
CYP2C9*2	R144C
CYP2C9*3	I359L

CYP2C9*4	I359T
CYP2C9*5	D360E
CYP2C9*7	Y358C
CYP2D6*1	wild-type
CYP2D6*2	R296C, S486T
CYP2D6*9	K281del
CYP2D6*10	P34S, S486T
CYP2D6*17	T107I, R296C, S486T
CYP3A4*1	wild-type
CYP3A4*2	S222P
CYP3A4*3	M445T
CYP3A4*4	I118V
CYP3A4*5	P218R
CYP3A4*15	R162Q

Table 5 Polymorphic forms of P450 2C9, 2D6 and 3A4 cloned

The following PCR primers were used.

```
CYP2C9*2F:  5'-TGTGTTCAAGAGGAAGCCCGCTG-3'
CYP2C9*2R:  5'-GTCCTCAATGCTGCTCTTCCCCATC-3'
CYP2C9*3F:  5'-CTTGACCTTCTCCCCACCAGCCTG-3'
CYP2C9*3R:  5'-GTATCTCTGGACCTCGTGCACCAC-3'
CYP2C9*4F:  5'-CTGACCTTCTCCCCACCAGCCTG-3'
CYP2C9*4R:  5'-TGTATCTCTGGACCTCGTGCAC-3'
CYP2C9*5F:  5'-GCTTCTCCCCACCAGCCTGC-3'
CYP2C9*5R:  5'-TCAATGTATCTCTGGACCTCGTGC-3'
CYP2C9*7F   5'-GCATTGACCTTCTCCCCACCAGC-3'
CYP2C9*7R:  5'-CACCACGTGCTCCAGGTCTCTA-3'
```

*CYP2D6*10AF1:* *5'-TATTCTCACTGGCCATTACGGCCGTGGACCTGATGCACCGGCGCCAACGCTGG GCTGCACGCTACTCACCAGGCCCCCTGC-3'*

*CYP2D6*10AR1:* *5'-GCGGGGCACAGCACAAAGCTCATAGGGGGATGGGCTCACCAGGAAAGCAAAG-3'*

*CYP2D6*17F: 5'-TCCAGATCCTGGGTTTCGGGC-3'*

*CYP2D6*17R: 5'-TGATGGGCACAGGCGGGCGGTC-3'*

*CYP2D6*9F: 5'-GCCAAGGGGAACCCTGAGAGC-3'*
CYP2D6*9R: 5'-CTCCATCTCTGCCAGGAAGGC-3'

*CYP3A4*2F: 5'-CCAATAACAGTCTTTCCATTCCTC-3'*

*CYP3A4*2R: 5'-GAGAAAGAATGGATCCAAAAAATC-3'*

*CYP3A4*3F: 5'-CGAGGTTTGCTCTCATGACCATG-3'*
CYP3A4*3R: 5'-TGCCAATGCAGTTTCTGGGTCCAC-3'
CYP3A4*4F: 5'-GTCTCTATAGCTGAGGATGAAG-3'
CYP3A4*4R: 5'-GGCACTTTTCATAAATCCCACTG-3'
CYP3A4*5F: 5'-GATTCTTTCTCTCAATAACAGTC-3'
CYP3A4*5R: 5'-GATCCAAAAAATCAAATCTTAAA-3'
CYP3A4*15F: 5'-AGGAAGCAGAGACAGGCAAGC-3'
CYP3A4*15R: 5'-GCCTCAGATTTCTCACCAACAC-3'

Example 6: Expression and Purification of P450 3A4

E. coli XL-10 gold (Stratagene) was used as a host for expression cultures of P450 3A4. Starter cultures were grown overnight in LB media supplemented with 100mg per litre ampicillin. 0.5 litre Terrific Broth media plus 100mg per litre ampicillin and 1mM thiamine and trace elements were inoculated with 1/100 dilution of the overnight starter cultures. The flasks were shaken at 37°C until cell density OD_{600} was 0.4 then δ-Aminolevulinic acid (ALA) was added to the cells at 0.5mM for 20 min at 30°C. The cells were supplemented with

50μM biotin then induced with optimum concentration of IPTG (30- 100μM) then shaken overnight at 30°C.

The E. coli cells from 0.5 litre cultures were divided into 50 ml aliquots, cells pelleted by centrifugation and cell pellets stored at –20°C. Cells from each pellet were lysed by resuspending in 5ml buffer A (100mM Tris buffer pH 8.0 containing 100 mM EDTA, 10mM β-mercaptoethanol, 10x stock of Protease inhibitor cocktail- Roche 1836170, 0.2mg/ml Lysozyme). After 15 minutes incubation on ice 40 ml of ice-cold deionised water was added to each resuspended cell pellet and mixed. 20 mM Magnesium Chloride and 5μg/ml DNaseI were added. The cells were incubated for 30 min on ice with gentle shaking after which the lysed E.Coli cells were pelletted by centrifugation for 30 min at 4000 rpm. The cell pellets were washed by resuspending in 10 ml buffer B (100mM Tris buffer pH 8.0 containing 10mM β-mercaptoethanol and a 10x stock of Protease inhibitor cocktail- Roche 1836170) followed by centrifugation at 4000 rpm. Membrane associated protein was then solubilised by the addition of 2 ml buffer C (50mM potassium phosphate pH 7.4, 10x stock of Protease inhibitor cocktail- Roche 1836170, 10 mM β-mercaptoethanol, 0.5 M NaCl and 0.3% (v/v) Igepal CA-630) and incubating on ice with gentle agitation for 30 minutes before centrifugation at 10,000g for 15 min at 4°C and the supernatant (Fig. 14) was then applied to Talon resin (Clontech).

A 0.5 ml column of Ni-NTA agarose (Qiagen) was poured in disposable gravity columns and equilibrated with 5 column volumes of buffer C. Supernatant was applied to the column after which the column was successively washed with 4 column volumes of buffer C, 4 column volumes of buffer D (50mM potassium phosphate pH 7.4, 10x stock of Protease inhibitor cocktail- Roche 1836170, 10 mM β-mercaptoethanol, 0.5 M NaCl and 20% (v/v) Glycerol) and 4 column

volumes of buffer D + 50 mM Imidazole before elution in 4 column volumes of buffer D + 200 mM Imidazole (Fig. 15). 0.5ml fractions were collected and protein containing fractions were pooled aliquoted and stored at -80°C.

Example 7: Determination of heme incorporation into P450s

Purified P450s were diluted to a concentration of 0.2 mg / ml in 20 mM potassium phosphate (pH 7.4) in the presence and absence of 10 mM KCN and an absorbance scan measured from 600 – 260 nm. The percentage bound heme was calculated based on an extinction coefficient ε_{420} of 100 $mM^{-1}cm^{-1}$.

Example 8: Reconstitution and assay of cytochrome P450 enzymes into liposomes with NADPH-cytochrome P450 reductase

Liposomes are prepared by dissolving a 1:1:1 mixture of 1,2-dilauroyl-sn-glycero-3-phosphocholine, 1,2-dileoyl-sn-glycero-3-phosphocholine, 1,2-dilauroyl-sn-glycero-3-phosphoserine in chloroform, evaporating to dryness and subsequently resuspending in 20 mM potassium phosphate pH 7.4 at 10 mg/ml. 4 μg of liposomes are added to a mixture of purified P450 2D6 (20 pmol), NADPH P450 reductase (40 pmol), cytochrome b5 (20 pmol) in a total volume of 10 μl and preincubated for 10 minutes at 37°C.

After reconstitution of cytochrome P450 enzymes into liposomes, the liposomes are diluted to 100 μl in assay buffer in a black 96 well plate, containing HEPES / KOH (pH 7.4, 50 mM), NADP+ (2.6 mM), glucose-6-phosphate (6.6 mM), $MgCl_2$ (6.6 mM) and glucose-6-phosphate dehyrogenase (0.4 units / ml). Assay buffer also contains an appropriate fluorogenic substrate for the cytochrome P450 isoform to be assayed: for P450 2D6 AMMC, for P450 3A4 dibenzyl fluorescein (DBF) or resorufin benzyl ether (BzRes) can be used and for 2C9 dibenzyl fluorescein (DBF). The reactions are stopped by the addition of 'stopping solution' (80% acetonitrile buffered with Tris) and products are read

using the appropriate wavelength filter sets in a fluorescent plate reader (Fig. 16).

P450s can also be activated chemically by, for example, the addition of 200 μM cumene hydroperoxide in place of the both the co-enzymes and regeneration solution (Fig. 17).

In addition fluorescently measured rates of turnover can be measured in the presence of inhibitors.

Example 9: Detection of Drug Binding to immobilised P450s CYP3A4

Purified CYP3A4 (10μg/ml in 50mM HEPES/0.01% CHAPS, pH 7.4) was placed in streptavidin immobiliser plates (Exiqon) (100μl per well) and shaken on ice for 1 hour. The wells were aspirated and washed twice with 50mM HEPES/0.01% CHAPS. [^{3}H]-ketoconazole binding to immobilised protein was determined directly by scintillation counting. Saturation experiments were performed using [^{3}H]ketoconazole (5Ci/mmol, American Radiochemicals Inc., St. Louis) in 50mM HEPES pH 7.4, 0.01% CHAPS and 10% Superblock (Pierce) (Figure 18). Six concentrations of ligand were used in the binding assay (25 – 1000nM) in a final assay volume of 100μl. Specific binding was defined as that displaced by 100μM ketoconazole. Each measurement was made in duplicate. After incubation for 1 hour at room temperature, the contents of the wells were aspirated and the wells washed three times with 150μl ice cold assay buffer. 100μl MicroScint 20 (Packard) was added to each well and the plates counted in a Packard TopCount microplate scintillation counter (Fig. 18).

Example 10 Chemical activation of tagged, immobilised CYP3A4

CYP3A4 was immobilised in streptavidin immobiliser plates as described in Example 9 and was then incubated with dibenzyl fluorescein and varying concentrations (0-300μM) of cumene hydrogen peroxide. End point assays demonstrated that the tagged, immobilised CYP3A4 was functional in a turn-over assay with chemical activation (Fig. 19).

Example 11: Immobilisation of P450s through gel encapsulation of liposomes or microsomes

After reconstitution of cytochrome P450 enzymes together with NADPH-cytochrome P450 reductase in liposomes or microsomes, these can then be immobilised on to a surface by encapsulation within a gel matrix such as agarose, polyurethane or polyacrylamide.

For example, low melting temperature (LMT) (1% w/v) agarose was dissolved in 200mM potassium phosphate pH 7.4. This was then cooled to 37 °C on a heating block. Microsomes containing cytochrome P450 3A4, cytochrome b5 and NADPH-cytochrome P450 reductase were then diluted into the LMT agarose such that 50 μl of agarose contained 20, 40 and 20 pmol of P450 3A4, NADPH-cytochrome P450 reductase and cytochrome b5 respectively. 50 μl of agarose-microsomes was then added to each well of a black 96 well microtitre plate and allowed to solidify at room temperature.

To each well, 100 μl of assay buffer was added and the assay was conducted as described previously (for example, Example 8) for conventional reconstitution assay. From the data generated a comparison of the fundamental kinetics of

BzRes oxidation and ketoconazole inhibition was made (Table 6) which showed that the activity of the CYP3A4 was retained after gel-encapsulation.

	Gel encapsulated	Soluble
BzRes Oxidation		
K_M (μM)	49 (18)	20 (5)
V_{max} (% of soluble)	50 (6)	100 (6)
Ketoconazole inhibition		
IC50 (nM)	86 (12)	207 (54)

Table 6 Comparison of kinetic parameters for Bz Rez oxidation and inhibition by ketoconazole for cytochrome P450 3A4 microsomes in solution and encapsulated in agarose. For estimation of K_M and V_{max} for BzRes assays were performed in the presence of varying concentrations of BzRes up to 320 μM. Ketoconazole inhibition was performed at 50 μM BzRes with 7 three-fold dilutions of ketoconazole from 5 μM. Values in parenthesis indicate standard errors derived from the curve fitting.

The activity of the immobilised P450s was assessed over a period of 7 days (Fig. 20). Aliquots of the same protein preparation stored under identical conditions, except that they were not gel-encapsulated, were also assayed over the same period, which revealed that the gel encapsualtion confers significant stability to the P450 activity.

Example 12: Quantitative determination of affect of 3A4 polymorphisms on activity

Purified cytochrome P450 3A4 isoforms *1, *2, *3, *4, *5 & *15 (approx 1 μg) were incubated in the presence of BzRes and cumene hydrogen peroxide (200

μM) in the absence and presence of ketoconazole at room temperature in 200 mM KPO_4 buffer pH 7.4 in a total volume of 100 μl in a 96 well black microtitre plate. A minimum of duplicates were performed for each concentration of BzRes or ketoconazole.

Resorufin formation of was measured over time by the increase in fluorescence (520 nm and 580 nm excitation and emission filters respectively) and initial rates were calculated from progress curves (Fig. 21).

For estimation of K_M^{app} and V_{max}^{app} for BzRes, background rates were first subtracted from the initial rates and then were plotted against BzRes concentration and curves were fitted describing conventional Michaelis-Menton kinetics:

$V=V_{max} / (1 + (K_M / S))$

where V and S are initial rate and substrate concentration respectively. V_{max} values were then normalised for cytochrome P450 concentration and scaled to the wild-type enzyme (Table 7).

For estimation of IC_{50} for ketoconazole, background rates were first subtracted from the initial rates which were then converted to a % of the uninhibited rate and plotted against ketoconazole concentration (Fig. 22). IC_{50} inhibition curves were fitted using the equation:

$V = 100 / (1 + (I / IC_{50}))$

where V and I are initial rate and inhibitor concentration respectively. The data obtained is shown in Table 7:

	V_{max} BzRes	K_M BzRes (μM)	IC_{50} ketoconazole (μM)
3A4*WT	100 (34)	104 (25)	0.91 (0.45)
3A4*2	65 (9)	62 (4)	0.44 (0.11)
3A4*3	93 (24)	54 (13)	1.13 (0.16)
3A4*4	69 (22)	111 (18)	0.88 (0.22)
3A4*5	59 (16)	101 (11)	1.96 (0.96)
3A4*15	111 (23)	89 (11)	0.59 (0.20)

Table 7 Kinetic parameters for BzRes turnover and its inhibition by ketoconazole for cytochrome P450 3A4 isoforms. The parameters were obtained from the fits of Michaelis-Menton and IC_{50} inhibition curves to the data in Figs. 21 & 22. Values in parenthesis are standard errors obtained from the curve fits.

Example 13: Array-based assay of immobilised CYP3A4 polymorphisms

Cytochrome P450 polymorphisms can be assayed in parallel using an array format to identify subtle differences in activity with specific small molecules. For example, purified cytochrome P450 3A4 isoforms *1, *2, *3, *4, *5 & *15 can be individually reconstituted in to liposomes with NADPH-cytochrome P450 reductase as described in Example 11. The resultant liposomes preparation can then be diluted into LMP agarose and immobilised into individual wells of a black 96 well microtitre plate as described in Example 11.

The immobilised proteins can then be assay ed as described in Example 11 by adding 100μl of assay buffer containing BzRes +/- ketoconazole to each well.

Chemical activation (as described in Example 12) can also be used in an array format. For example, purified cytochrome P450 3A4 isoforms *1, *2, *3, *4, *5 &

*15 can be individually reconstituted in to liposomes without NADPH-cytochrome P450 reductase and the resultant liposomes can be immobilised via encapsulation in agarose as described in Example 11. The cytochrome P450 activity in each well can then be measured as described in Example 12 by 100μl of 200 mM KPO_4 buffer pH 7.4 containing BzRes and cumene hydrogen peroxide (200 μM), +/- ketoconazole, to each well.

In summary, the Inventors have developed a novel protein array technology for massively parallel, high-throughout screening of SNPs for the biochemical activity of the encoded proteins. Its applicability was demonstrated through the analysis of various functions of wild type p53 and 46 SNP versions of p53 as well as with allelic variants of p450. The same surface and assay detection methodologies can now be applied to other more diverse arrays currently being developed. Due to the small size of the collection of proteins being studied here, the spot density of our arrays was relatively small, and each protein was spotted in quadruplicate. Using current robotic spotting capabilities it is possible to increase spot density to include over 10,000 proteins per array.

CLAIMS

1. A protein array comprising a surface upon which are deposited at spatially defined locations at least two protein moieties characterised in that said protein moieties are those of naturally occurring variants of a DNA sequence of interest.

2. A protein array as claimed in claim 1 wherein said variants map to the same chromosomal locus.

3. A protein array as claimed in claim 1 or 2 wherein the one or more protein moieties are derived from synthetic equivalents of naturally occurring variants of a DNA sequence of interest.

4. A protein array as claimed in claim 1 or claim 2 wherein said at least two protein moieties comprise a protein moiety expressed by a wild type gene of interest together with at least one protein moiety expressed by one or more genes containing one or more naturally occurring mutations thereof.

5. A protein array as claimed in claim 4 wherein said mutations are selected from the group consisting of, a mis-sense mutation, a single nucleotide polymorphism, a deletion mutation, and an insertion mutation.

6. A protein array as claimed in any of the preceding claims wherein the protein moieties comprise proteins associated with a disease state, drug metabolism or those which are uncharacterised.

7. A protein array as claimed in any of the preceding claims wherein the protein moieties encode wild type p53 and allelic variants thereof.

8. A protein array as claimed in any of the claims 1 to 6 wherein the protein moieties encode a drug metabolising enzyme.

9. A protein array as claimed in claim 8 wherein the drug metabolising enzyme is wild type p450 and allelic variants thereof.

10. A method of making a protein array comprising the steps of

a) providing DNA coding sequences which are those of two or more naturally occurring variants of a DNA sequence of interest

b) expressing said coding sequences to provide one or more individual protein moieties

c) purifying said protein moieties

d) depositing said protein moieties at spatially defined locations on a surface to give an array.

11. The method as claimed in claim 10, wherein steps c) and d) are combined in a single step by the simultaneous purification and isolation of the protein moieties on the array via an incorporated tag.

12. The method as claimed in claim 10, wherein step c) is omitted and said individual protein moieties are present with other proteins from an expression host cell.

13. The method as claimed in claim 10, wherein said DNA sequence of interest encodes a protein associated with a disease state, drug metabolism or is uncharacterised.

14. The method as claimed in claim 13, wherein said DNA sequence of interest encodes p53.

15. The method as claimed in claim 13, wherein said DNA sequence of interest encodes a drug metabolising enzyme.

16. The method as claimed in claim 15, wherein said drug metabolising enzyme is wild type p450 and allelic variants thereof.

17. Use of an array as claimed in any of claims 1 to 9 in the determination of the phenotype of a naturally occurring variant of a DNA sequence of interest wherein said DNA sequence is represented by at least one protein moiety derived therefrom and is present on said array.

18. A method of screening a set of protein moieties for molecules which interact with one or more proteins comprising the steps of

a) bringing one or more test molecules into contact with an array as claimed in any one of claims 1 to 9; which carries said set of protein moieties; and

b) detecting an interaction between one or more test molecules and one or more proteins on the array.

19. A method of simultaneously determining the relative properties of members of a set of protein moieties, comprising the steps of:

a) bringing an array as claimed in any one of claims 1 to 9 which carries said set of protein moieties into contact with one or more test substances, and

b) observing the interaction of said test substances with the set members on the array.

20. The method of claim 19 wherein one or more of said protein moieties are drug metabolising enzymes and wherein said enzymes are activated by contact with an accessory protein or by chemical treatment.

PCT/GB 2002 / 0 0 5 4 9 9

ABSTRACT

The Invention describe protein arrays and their use to assay, in a parallel fashion, the protein products of highly homologous or related DNA coding sequences.

By highly homologous or related it is meant those DNA coding sequences which share a common sequence and which differ only by one or more naturally occurring mutations such as single nucleotide polymorphisms, deletions or insertions, or those sequences which are considered to be haplotypes (a haplotype being a combination of variations or mutations on a chromosome, usually within the context of a particular gene). Such highly homologous or related DNA coding sequences are generally naturally occurring variants of the same gene. Arrays according to the invention have multiple for example, two or more, individual proteins deposited in a spatially defined pattern on a surface in a form whereby the properties, for example the activity or function of the proteins can be investigated or assayed in parallel by interrogation of the array.

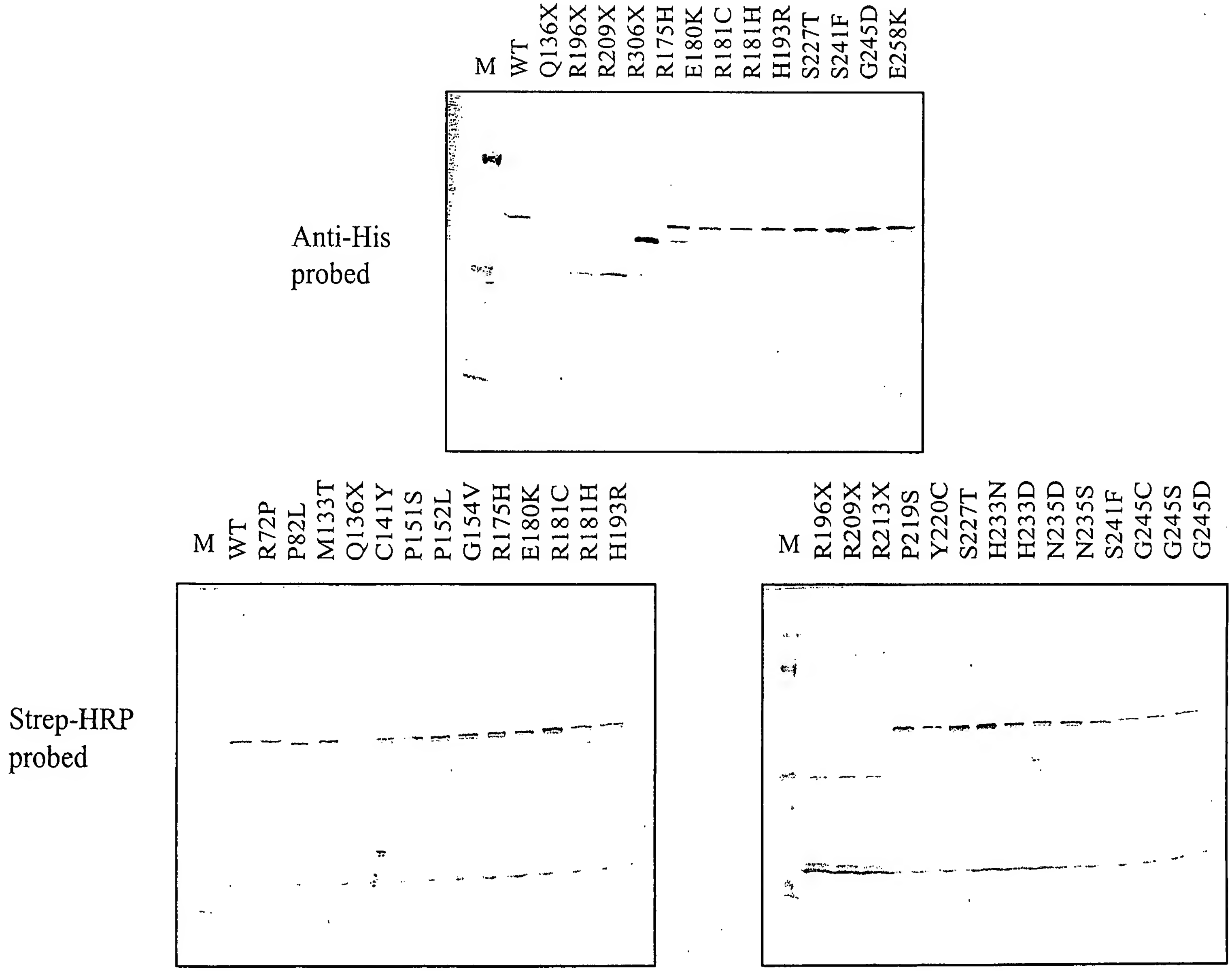


Figure 1

REPLACED

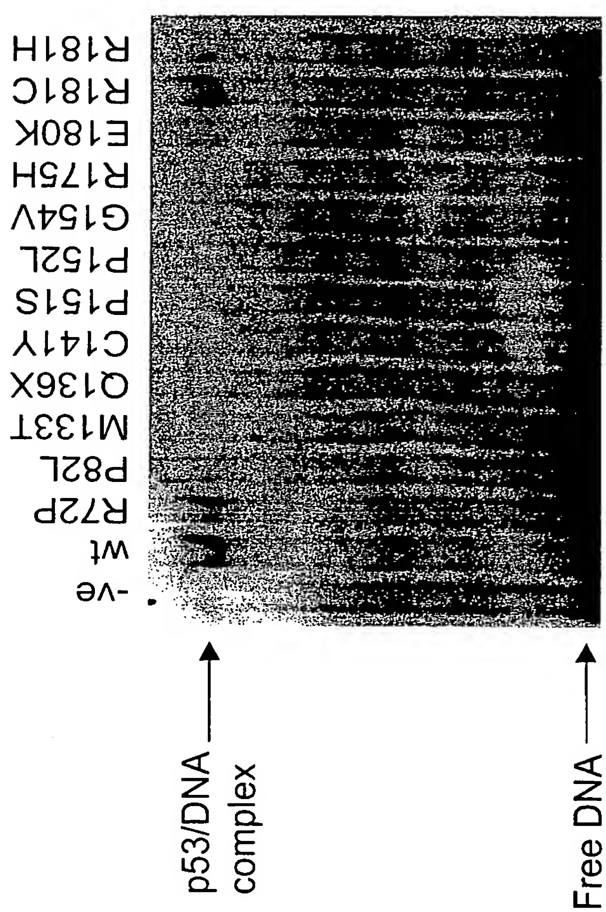
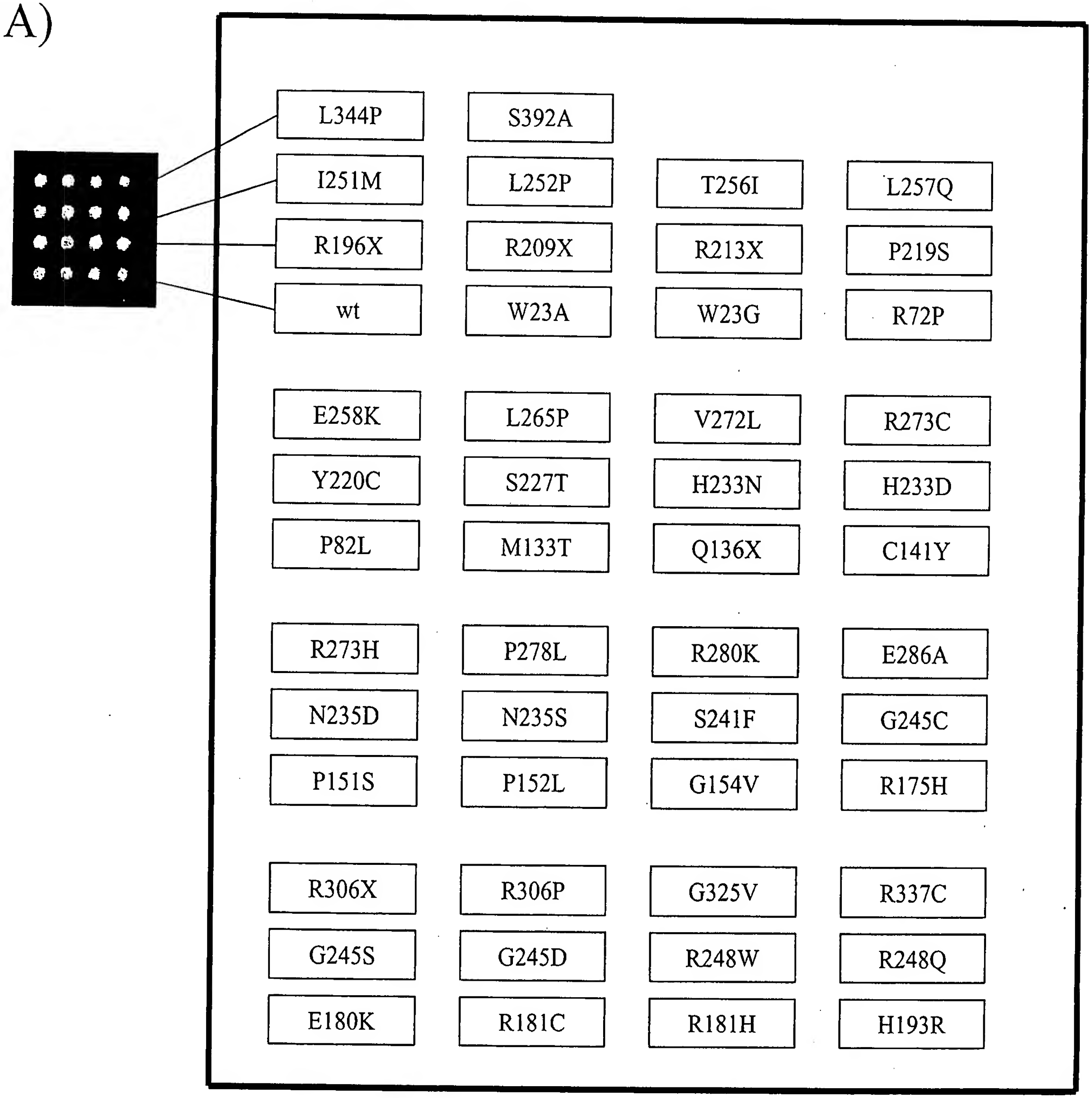

-ve
wt
R72P
P82L
M133T
Q136X
C141Y
P151S
P152L
G154V
R175H
E180K
R181C
R181H
p53/DNA complex
Free DNA

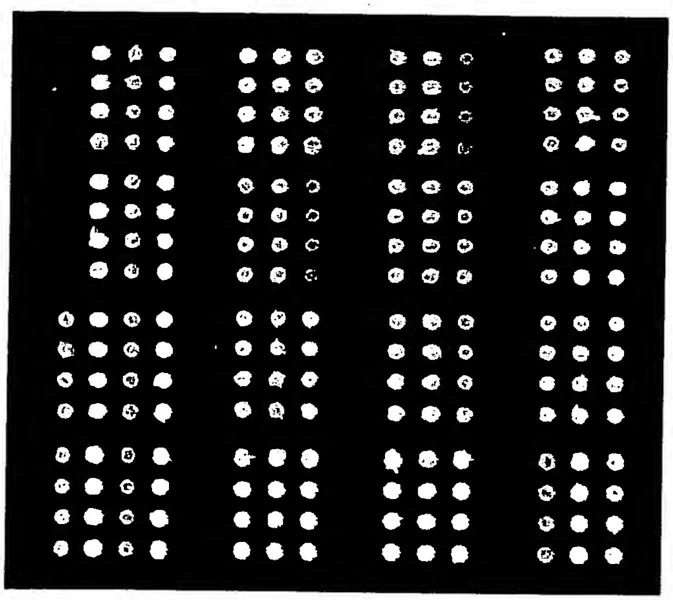
Figure 2

3A)



L344P
S392A
I251M
L252P
T256I
L257Q
R196X
R209X
R213X
P219S
wt
W23A
W23G
R72P
E258K
L265P
V272L
R273C
Y220C
S227T
H233N
H233D
P82L
M133T
Q136X
C141Y
R273H
P278L
R280K
E286A
N235D
N235S
S241F
G245C
P151S
P152L
G154V
R175H
R306X
R306P
G325V
R337C
G245S
G245D
R248W
R248Q
E180K
R181C
R181H
H193R

3B)



3C)

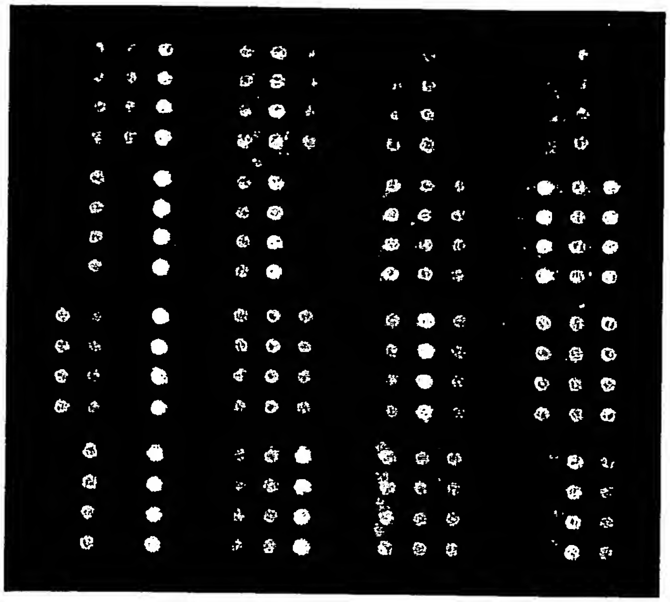


Figure 3

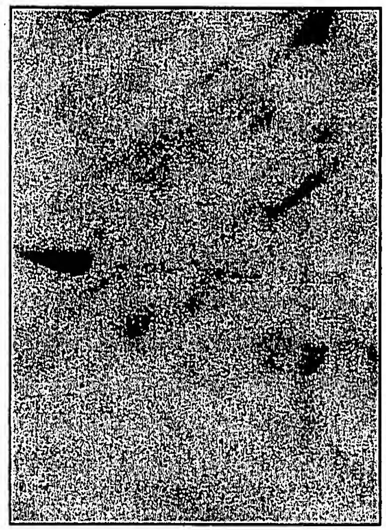


Figure 4

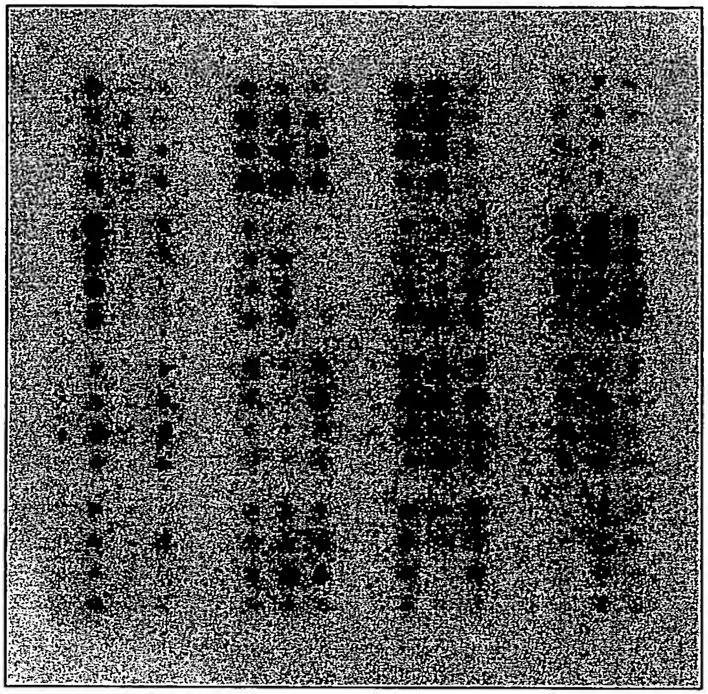


Figure 5

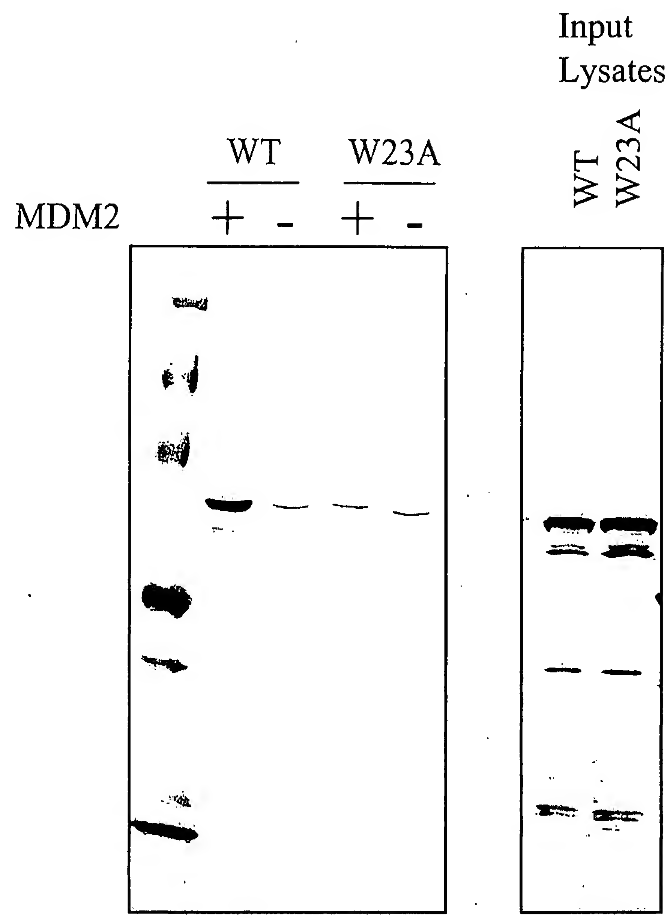


Figure 6

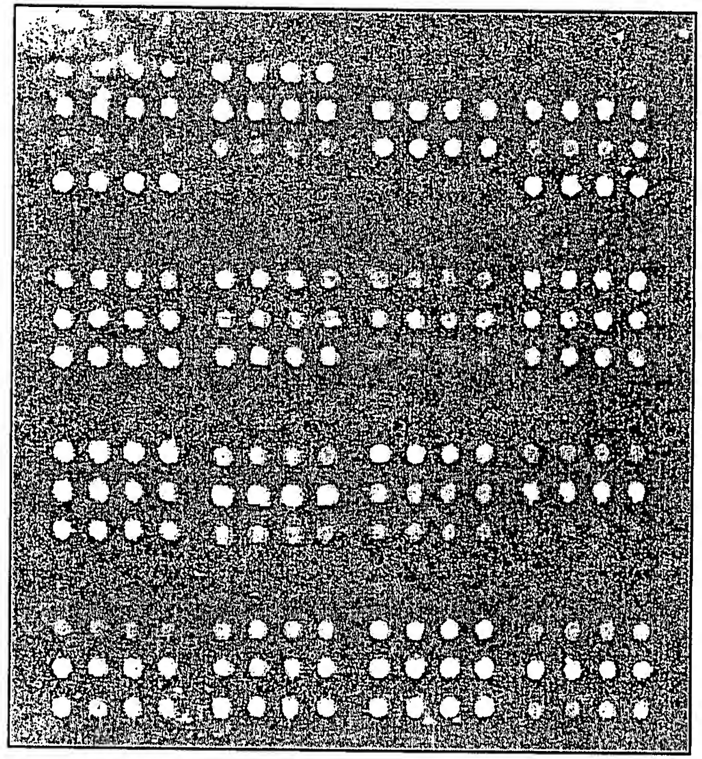


Figure 7

FIG 8b

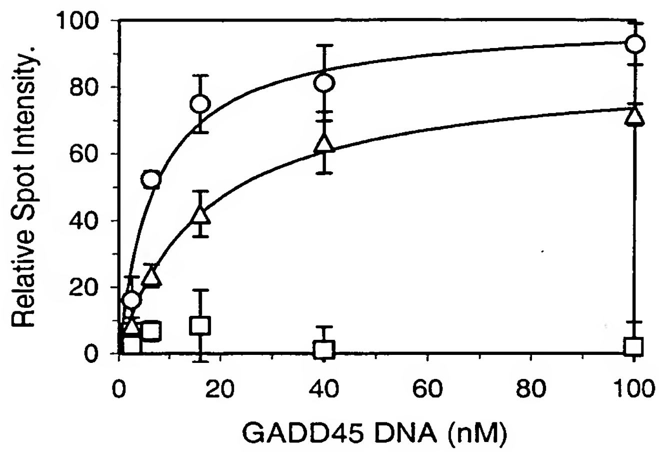
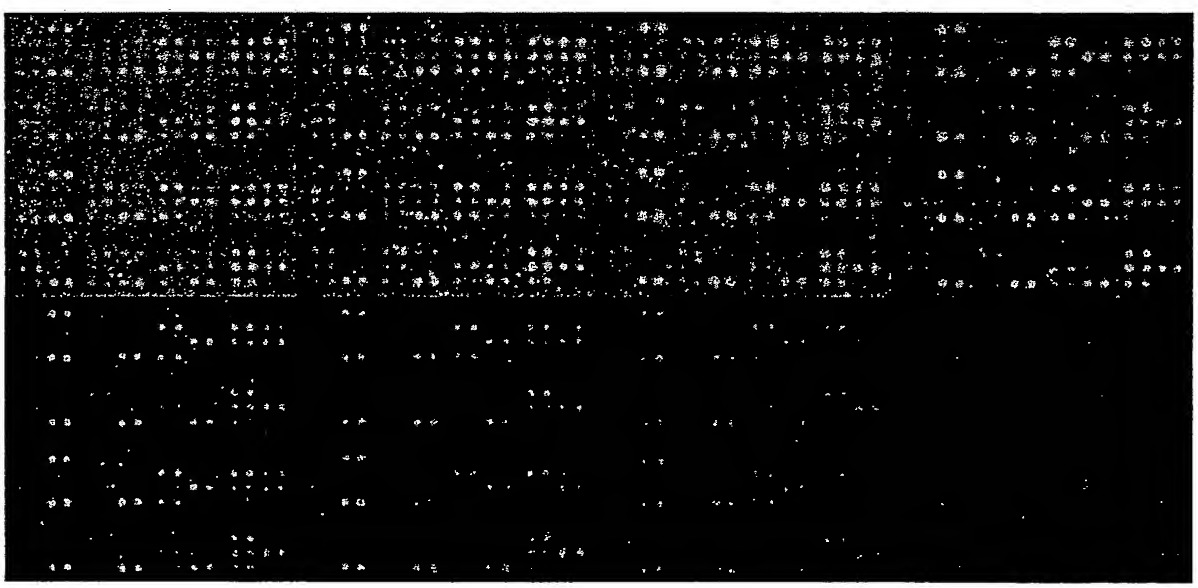


Fig 8a

100Nm

2.5 fold dilution



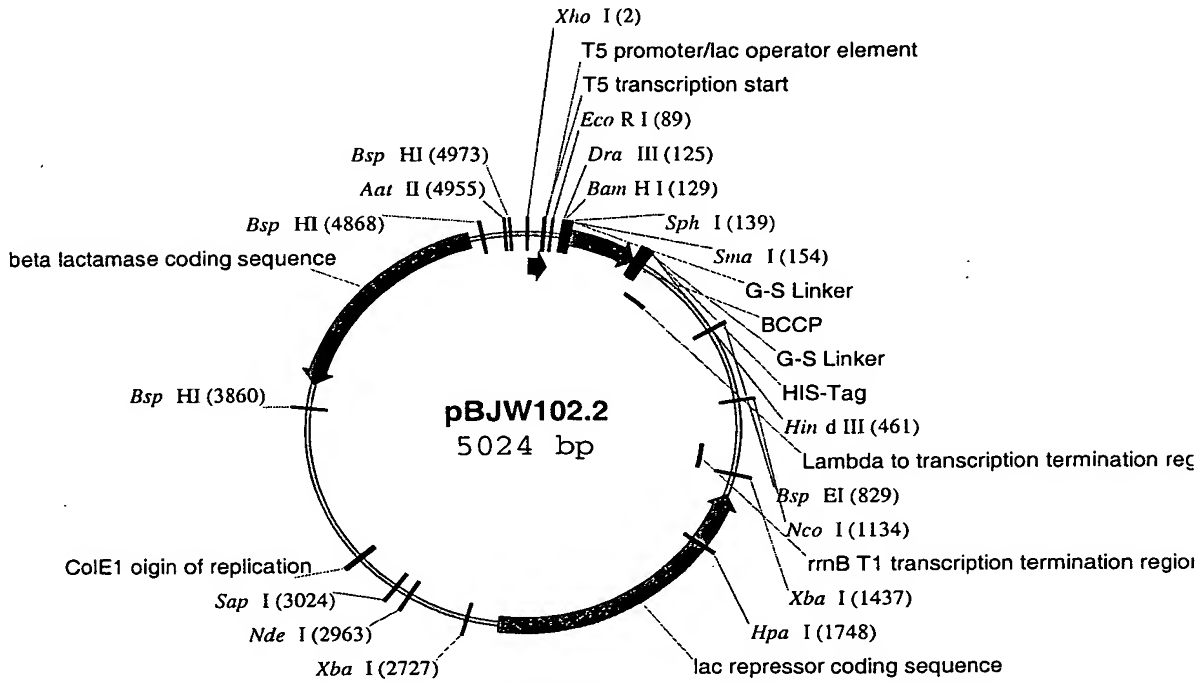

Xho I (2)
T5 promoter/lac operator element
T5 transcription start
Eco R I (89)
Dra III (125)
Bam H I (129)
Sph I (139)
Sma I (154)
G-S Linker
BCCP
G-S Linker
HIS-Tag
Hin d III (461)
Lambda to transcription termination reg
Bsp EI (829)
Nco I (1134)
rrnB T1 transcription termination regio
Xba I (1437)
Hpa I (1748)
lac repressor coding sequence
Xba I (2727)
Nde I (2963)
Sap I (3024)
ColE1 oigin of replication
Bsp HI (3860)
beta lactamase coding sequence
Bsp HI (4868)
Aat II (4955)
Bsp HI (4973)
pBJW102.2
5024 bp

Figure 9A

```
   1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA
  61 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGGCA
 121 CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGCGGTTC TGGCGCAGCA
 181 GCGGAAATCA GTGGTCACAT CGTACGTTCC CCGATGGTTG GTACTTTCTA CCGCACCCCA
 241 AGCCCGGACG CAAAAGCGTT CATCGAAGTG GGTCAGAAAG TCAACGTGGG CGATACCCTG
 301 TGCATCGTTG AAGCCATGAA AATGATGAAC CAGATCGAAG CGGACAAATC CGGTACCGTG
 361 AAAGCAATTC TGGTCGAAAG TGGACAACCG GTAGAATTTG ACGAGCCGCT GGTCGTCATC
 421 GAGGGTGGCA GCGGTTCTGG CCACCATCAC CATCACCATA AGCTTAATTA GCTGAGCTTG
 481 GACTCCTGTT GATAGATCCA GTAATGACCT CAGAACTCCA TCTGGATTTG TTCAGAACGC
 541 TCGGTTGCCG CCGGGCGTTT TTTATTGGTG AGAATCCAAG CTAGCTTGGC GAGATTTTCA
 601 GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA ATCACTGGAT ATACCACCGT TGATATATCC
 661 CAATGGCATC GTAAAGAACA TTTTGAGGCA TTTCAGTCAG TTGCTCAATG TACCTATAAC
 721 CAGACCGTTC AGCTGGATAT TACGGCCTTT TTAAAGACCG TAAAGAAAAA TAAGCACAAG
 781 TTTTATCCGG CCTTTATTCA CATTCTTGCC CGCCTGATGA ATGCTCATCC GGAATTTCGT
 841 ATGGCAATGA AAGACGGTGA GCTGGTGATA TGGGATAGTG TTCACCCTTG TTACACCGTT
 901 TTCCATGAGC AAACTGAAAC GTTTTCATCG CTCTGGAGTG AATACCACGA CGATTTCCGG
 961 CAGTTTCTAC ACATATATTC GCAAGATGTG GCGTGTTACG GTGAAAACCT GGCCTATTTC
1021 CCTAAAGGGT TTATTGAGAA TATGTTTTTC GTCTCAGCCA ATCCCTGGGT GAGTTTCACC
1081 AGTTTTGATT TAAACGTGGC CAATATGGAC AACTTCTTCG CCCCCGTTTT CACCATGGGC
1141 AAATATTATA CGCAAGGCGA CAAGGTGCTG ATGCCGCTGG CGATTCAGGT TCATCATGCC
1201 GTTTGTGATG GCTTCCATGT CGGCAGAATG CTTAATGAAT TACAACAGTA CTGCGATGAG
1261 TGGCAGGGCG GGGCGTAATT TTTTTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGGGTA
1321 ATGACTCTCT AGCTTGAGGC ATCAAATAAA ACGAAAGGCT CAGTCGAAAG ACTGGGCCTT
1381 TCGTTTTATC TGTTGTTTGT CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCCTCTAG
1441 ATTACGTGCA GTCGATGATA AGCTGTCAAA CATGAGAATT GTGCCTAATG AGTGAGCTAA
1501 CTTACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT CGGGAAACCT GTCGTGCCAG
1561 CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCCAGGGT
1621 GGTTTTTCTT TTCACCAGTG AGACGGGCAA CAGCTGATTG CCCTTCACCG CCTGGCCCTG
1681 AGAGAGTTGC AGCAAGCGGT CCACGCTGGT TTGCCCCAGC AGGCGAAAAT CCTGTTTGAT
1741 GGTGGTTAAC GGCGGGATAT AACATGAGCT GTCTTCGGTA TCGTCGTATC CCACTACCGA
1801 GATATCCGCA CCAACGCGCA GCCCGGACTC GGTAATGGCG CGCATTGCGC CCAGCGCCAT
1861 CTGATCGTTG GCAACCAGCA TCGCAGTGGG AACGATGCCC TCATTCAGCA TTTGCATGGT
1921 TTGTTGAAAA CCGGACATGG CACTCCAGTC GCCTTCCCGT TCCGCTATCG GCTGAATTTG
1981 ATTGCGAGTG AGATATTTAT GCCAGCCAGC CAGACGCAGA CGCGCCGAGA CAGAACTTAA
2041 TGGGCCCGCT AACAGCGCGA TTTGCTGGTG ACCCAATGCG ACCAGATGCT CCACGCCCAG
2101 TCGCGTACCG TCTTCATGGG AGAAAATAAT ACTGTTGATG GGTGTCTGGT CAGAGACATC
2161 AAGAAATAAC GCCGGAACAT TAGTGCAGGC AGCTTCCACA GCAATGGCAT CCTGGTCATC
2221 CAGCGGATAG TTAATGATCA GCCCACTGAC GCGTTGCGCG AGAAGATTGT GCACCGCCGC
2281 TTTACAGGCT TCGACGCCGC TTCGTTCTAC CATCGACACC ACCACGCTGG CACCCAGTTG
2341 ATCGGCGCGA GATTTAATCG CCGCGACAAT TTGCGACGGC GCGTGCAGGG CCAGACTGGA
2401 GGTGGCAACG CCAATCAGCA ACGACTGTTT GCCCGCCAGT TGTTGTGCCA CGCGGTTGGG
2461 AATGTAATTC AGCTCCGCCA TCGCCGCTTC CACTTTTTCC CGCGTTTTCG CAGAAACGTG
2521 GCTGGCCTGG TTCACCACGC GGGAAACGGT CTGATAAGAG ACACCGGCAT ACTCTGCGAC
2581 ATCGTATAAC GTTACTGGTT TCACATTCAC CACCCTGAAT TGACTCTCTT CCGGGCGCTA
2641 TCATGCCATA CCGCGAAAGG TTTTGCACCA TTCGATGGTG TCGGAATTTC GGGCAGCGTT
2701 GGGTCCTGGC CACGGGTGCG CATGATCTAG AGCTGCCTCG CGCGTTTCGG TGATGACGGT
2761 GAAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG CTTGTCTGTA AGCGGATGCC
2821 GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA GCGGGTGTTG GCGGGTGTCG GGGCGCAGCC
2881 ATGACCCAGT CACGTAGCGA TAGCGGAGTG TATACTGGCT TAACTATGCG GCATCAGAGC
2941 AGATTGTACT GAGAGTGCAC CATATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA
3001 AATACCGCAT CAGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC
3061 GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG
3121 GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
3181 AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
3241 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC
3301 CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG
3361 CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT
3421 CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC
3481 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
3541 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
3601 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
3661 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
3721 CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG
3781 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
3841 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA
```

```
3901 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
3961 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG
4021 TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
4081 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
4141 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
4201 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
4261 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
4321 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
4381 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
4441 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
4501 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
4561 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
4621 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
4681 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
4741 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
4801 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
4861 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC
4921 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
4981 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT TCAC
```

Figure 9B

```
           Dra III              Sph I                 Sma I
115  ATGGCA CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGC
     TACCGT GAATCACCCT AGGCGTACGC TCGAGCCATG GGGCCCCCAC CGTCG
```

Figure 9C

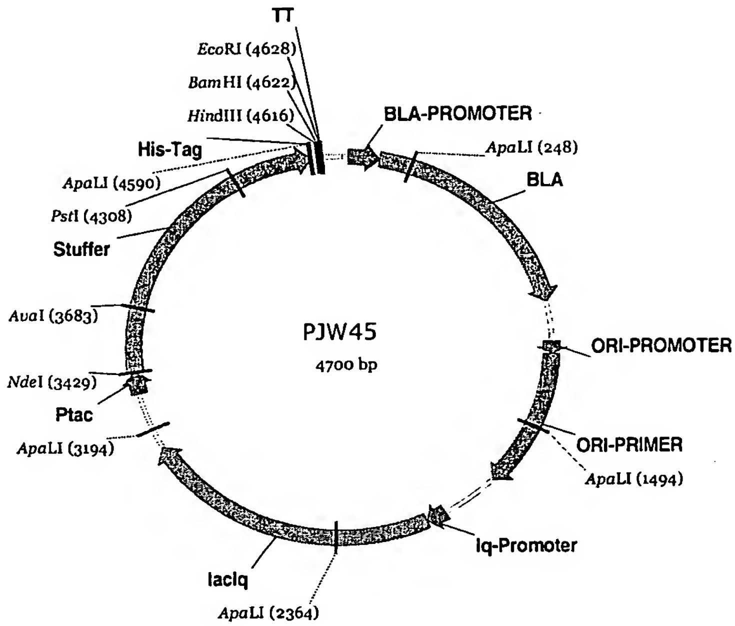

TT
EcoRI (4628)
BamHI (4622)
HindIII (4616)
His-Tag
ApaLI (4590)
PstI (4308)
Stuffer
AvaI (3683)
NdeI (3429)
Ptac
ApaLI (3194)
BLA-PROMOTER
ApaLI (248)
BLA
PJW45
4700 bp
ORI-PROMOTER
ORI-PRIMER
ApaLI (1494)
Iq-Promoter
lacIq
ApaLI (2364)

Figure 10A

PCT/GB 2002 / 0 0 5 4 9 9

```
   1 CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
  61 ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA
 121 AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT
 181 TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
 241 AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
 301 GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG
 361 CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
 421 AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG
 481 TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
 541 TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG
 601 TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG
 661 ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
 721 TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC
 781 CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
 841 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG
 901 TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
 961 AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC
1021 TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG
1081 ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
1141 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC
1201 AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
1261 TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT
1321 AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC
1381 TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
1441 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC
1501 AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCATTGAG
1561 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG
1621 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
1681 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
1741 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT
1801 TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
1861 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG
1921 AGGAAGCCCA GGACCCAACG CTGCCCGAAA TTCCGACACC ATCGAATGGT GCAAAACCTT
1981 TCGCGGTATG GCATGATAGC GCCCGGAAGA GAGTCAATTC AGGGTGGTGA ATGTGAAACC
2041 AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCG TTTCCCGCGT
2101 GGTGAACCAG GCCAGCCACG TTTCTGCGAA AACGCGGGAA AAAGTGGAAG CGGCGATGGC
2161 GGAGCTGAAT TACATTCCCA ACCGCGTGGC ACAACAACTG GCGGGCAAAC AGTCGTTGCT
2221 GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACGCGCCG TCGCAAATTG TCGCGGCGAT
2281 TAAATCTCGC GCCGATCAAC TGGGTGCCAG CGTGGTGGTG TCGATGGTAG AACGAAGCGG
2341 CGTCGAAGCC TGTAAAGCGG CGGTGCACAA TCTTCTCGCG CAACGCGTCA GTGGGCTGAT
2401 CATTAACTAT CCGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT GCACTAATGT
2461 TCCGGCGTTA TTTCTTGATG TCTCTGACCA GACACCCATC AACAGTATTA TTTTCTCCCA
2521 TGAAGACGGT ACGCGACTGG GCGTGGAGCA TCTGGTCGCA TTGGGTCACC AGCAAATCGC
```

2581 GCTGTTAGCG GGCCCATTAA GTTCTGTCTC GGCGCGTCTG CGTCTGGCTG GCTGGCATAA
2641 ATATCTCACT CGCAATCAAA TTCAGCCGAT AGCGGAACGG GAAGGCGACT GGAGTGCCAT
2701 GTCCGGTTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCCA CTGCGATGCT
2761 GGTTGCCAAC GATCAGATGG CGCTGGGCGC AATGCGCGCC ATTACCGAGT CCGGGCTGCG
2821 CGTTGGTGCG GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT CATGTTATAT
2881 CCCGCCGTTA ACCACCATCA AACAGGATTT TCGCCTGCTG GGGCAAACCA GCGTGGACCG
2941 CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC CCGTCTCACT
3001 GGTGAAAAGA AAAACCACCC TGGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC
3061 CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA
3121 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACAATT CTCATGTTTG ACAGCTTATC
3181 ATCGACTGCA CGGTGCACCA ATGCTTCTGG CGTCAGGCAG CCATCGGAAG CTGTGGTATG
3241 GCTGTGCAGG TCGTAAATCA CTGCATAATT CGTGTCGCTC AAGGCGCACT CCCGTTCTGG
3301 ATAATGTTTT TTGCGCCGAC ATCATAACGG TTCTGGCAAA TATTCTGAAA TGAGCTGTTG
3361 ACAATTAATC ATCGGCTCGT ATAATGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG
3421 GAAACACATA TGAACGACTT TCATCGCGAT ACGTGGGCGG AAGTGGATTT GGACGCCATT
3481 TACGACAATG TGGCGAATTT GCGCCGTTTG CTGCCGGACG ACACGCACAT TATGGCGGTC
3541 GTGAAGGCGA ACGCCTATGG ACATGGGGAT GTGCAGGTGG CAAGGACAGC GCTCGAAGCG
3601 GGGGCCTCCC GCCTGGCGGT TGCCTTTTTG GATGAGGCGC TCGCTTTAAG GGAAAAAGGA
3661 ATCGAAGCGC CGATTCTAGT TCTCGGGGCT TCCCGTCCAG CTGATGCGGC GCTGGCCGCC
3721 CAGCAGCGCA TTGCCCTGAC CGTGTTCCGC TCCGACTGGT TGGAAGAAGC GTCCGCCCTT
3781 TACAGCGGCC CTATTCCTAT TCATTTCCAT TTGAAAATGG ACACCGGCAT GGGACGGCTT
3841 GGAGTGAAAG ACGAGGAGGA GACGAAACGA ATCGCAGCGC TGATTGAGCG CCATCCGCAT
3901 TTTGTGCTTG AAGGGGCGTA CACGCATTTT GCGACTGCGG ATGAGGTGAA CACCGATTAT
3961 TTTTCCTATC AGTATACCCG TTTTTTGCAC ATGCTCGAAT GGCTGCCGTC GCGCCCGCCG
4021 CTCGTCCATT GCGCCAACAG CGCAGCGTCG CTCCGTTTCC CTGACCGGAC GTTCAATATG
4081 GTCCGCTTCG GCATTGCCAT GTATGGGCTT GCCCCGTCGC CCGGCATCAA GCCGCTGCTG
4141 CCGTATCCAT TAAAAGAAGC ATTTTCGCTC CATAGCCGCC TCGTACACGT CAAAAAACTG
4201 CAACCAGGCG AAAAGGTGAG CTATGGTGCG ACGTACACTG CGCAGACGGA GGAGTGGATC
4261 GGGACGATTC CGATCGGCTA TGCGGACGGC TGGCTCCGCC GCCTGCAGCA CTTTCATGTC
4321 CTTGTTGACG GACAAAAGGC GCCGATTGTC GGCCGCATTT GCATGGACCA GTGCATGATC
4381 CGCCTGCCTG GGCCGCTGCC GGTCGGCACG AAGGTGACAC TGATTGGTCG CCAGGGGGAC
4441 GAGGTAATTT CCATTGATGA TGTCGCTCGC CATTTGGAAA CGATCAACTA CGAAGTGCCT
4501 TGCACGATCA GCTATCGAGT GCCCCGTATT TTTTTCCGCC ATAAGCGTAT AATGGAAGTG
4561 AGAAACGCCA TTGGCCGCGG GGAAAGCAGT GCACATCACC ATCACCATCA CTAAAAGCTT
4621 GGATCCGAAT TCAGCCCGCC TAATGAGCGG GCTTTTTTTT GAACAAAATT AGCTTGGCTG
4681 TTTTGGCGGA TGAGAGAAGA

Figure 10B

```
1     ATGGCTCTCA TCCCAGACTT GGCCATGGAA ACCTGGCTTC TCCTGGCTGT CAGCCTGGTG
61    CTCCTCTATC TATATGGAAC CCATTCACAT GGACTTTTTA AGAAGCTTGG AATTCCAGGG
121   CCCACACCTC TGCCTTTTTT GGGAAATATT TTGTCCTACC ATAAGGGCTT TTGTATGTTT
181   GACATGGAAT GTCATAAAAA GTATGGAAAA GTGTGGGGCT TTTATGATGG TCAACAGCCT
241   GTGCTGGCTA TCACAGATCC TGACATGATC AAAACAGTGC TAGTGAAAGA ATGTTATTCT
301   GTCTTCACAA ACCGGAGGCC TTTTGGTCCA GTGGGATTTA TGAAAAGTGC CATCTCTATA
361   GCTGAGGATG AAGAATGGAA GAGATTACGA TCATTGCTGT CTCCAACCTT CACCAGTGGA
421   AAACTCAAGG AGATGGTCCC TATCATTGCC CAGTATGGAG ATGTGTTGGT GAGAAATCTG
481   AGGCGGGAAG CAGAGACAGG CAAGCCTGTC ACCTTGAAAG ACGTCTTTGG GGCCTACAGC
541   ATGGATGTGA TCACTAGCAC ATCATTTGGA GTGAACATCG ACTCTCTCAA CAATCCACAA
601   GACCCCTTTG TGGAAAACAC CAAGAAGCTT TTAAGATTTG ATTTTTTGGA TCCATTCTTT
661   CTCTCAATAA CAGTCTTTCC ATTCCTCATC CCAATTCTTG AAGTATTAAA TATCTGTGTG
721   TTTCCAAGAG AAGTTACAAA TTTTTTAAGA AAATCTGTAA AAAGGATGAA AGAAAGTCGC
781   CTCGAAGATA CACAAAAGCA CCGAGTGGAT TTCCTTCAGC TGATGATTGA CTCTCAGAAT
841   TCAAAAGAAA CTGAGTCCCA CAAAGCTCTG TCCGATCTGG AGCTCGTGGC CCAATCAATT
901   ATCTTTATTT TTGCTGGCTA TGAAACCACG AGCAGTGTTC TCTCCTTCAT TATGTATGAA
961   CTGGCCACTC ACCCTGATGT CCAGCAGAAA CTGCAGGAGG AAATTGATGC AGTTTTACCC
1021  AATAAGGCAC CACCCACCTA TGATACTGTG CTACAGATGG AGTATCTTGA CATGGTGGTG
1081  AATGAAACGC TCAGATTATT CCCAATTGCT ATGAGACTTG AGAGGGTCTG CAAAAAAGAT
1141  GTTGAGATCA ATGGGATGTT CATTCCCAAA GGGGTGGTGG TGATGATTCC AAGCTATGCT
1201  CTTCACCGTG ACCCAAAGTA CTGGACAGAG CCTGAGAAGT TCCTCCCTGA AAGATTCAGC
1261  AAGAAGAACA AGGACAACAT AGATCCTTAC ATATACACAC CCTTTGGAAG TGGACCCAGA
1321  AACTGCATTG GCATGAGGTT TGCTCTCATG AACATGAAAC TTGCTCTAAT CAGAGTCCTT
1381  CAGAACTTCT CCTTCAAACC TTGTAAAGAA ACACAGATCC CCCTGAAATT AAGCTTAGGA
1441  GGACTTCTTC AACCAGAAAA ACCCGTTGTT CTAAAGGTTG AGTCAAGGGA TGGCACCGTA
1501  AGTGGAGCCT GA
```

Figure 11A

```
1     MALIPDLAME TWLLLAVSLV LLYLYGTHSH GLFKKLGIPG PTPLPFLGNI LSYHKGFCMF
61    DMECHKKYGK VWGFYDGQQP VLAITDPDMI KTVLVKECYS VFTNRRPFGP VGFMKSAISI
121   AEDEEWKRLR SLLSPTFTSG KLKEMVPIIA QYGDVLVRNL RREAETGKPV TLKDVFGAYS
181   MDVITSTSFG VNIDSLNNPQ DPFVENTKKL LRFDFLDPFF LSITVFPFLI PILEVLNICV
241   FPREVTNFLR KSVKRMKESR LEDTQKHRVD FLQLMIDSQN SKETESHKAL SDLELVAQSI
301   IFIFAGYETT SSVLSFIMYE LATHPDVQQK LQEEIDAVLP NKAPPTYDTV LQMEYLDMVV
361   NETLRLFPIA MRLERVCKKD VEINGMFIPK GVVVMIPSYA LHRDPKYWTE PEKFLPERFS
421   KKNKDNIDPY IYTPFGSGPR NCIGMRFALM NMKLALIRVL QNFSFKPCKE TQIPLKLSLG
481   GLLQPEKPVV LKVESRDGTV SGA*
```

Figure 11B

```
   1 ATGGATTCTC TTGTGGTCCT TGTGCTCTGT CTCTCATGTT TGCTTCTCCT TTCACTCTGG
  61 AGACAGAGCT CTGGGAGAGG AAAACTCCCT CCTGGCCCCA CTCCTCTCCC AGTGATTGGA
 121 AATATCCTAC AGATAGGTAT TAAGGACATC AGCAAATCCT TAACCAATCT CTCAAAGGTC
 181 TATGGCCCGG TGTTCACTCT GTATTTTGGC CTGAAACCCA TAGTGGTGCT GCATGGATAT
 241 GAAGCAGTGA AGGAAGCCCT GATTGATCTT GGAGAGGAGT TTTCTGGAAG AGGCATTTTC
 301 CCACTGGCTG AAAGAGCTAA CAGAGGATTT GGAATTGTTT TCAGCAATGG AAAGAAATGG
 361 AAGGAGATCC GGCGTTTCTC CCTCATGACG CTGCGGAATT TTGGGATGGG GAAGAGGAGC
 421 ATTGAGGACC GTGTTCAAGA GGAAGCCCGC TGCCTTGTGG AGGAGTTGAG AAAAACCAAG
 481 GCCTCACCCT GTGATCCCAC TTTCATCCTG GGCTGTGCTC CCTGCAATGT GATCTGCTCC
 541 ATTATTTTCC ATAAACGTTT TGATTATAAA GATCAGCAAT TTCTTAACTT AATGGAAAAG
 601 TTGAATGAAA ACATCAAGAT TTTGAGCAGC CCCTGGATCC AGATCTGCAA TAATTTTTCT
 661 CCTATCATTG ATTACTTCCC GGGAACTCAC AACAAATTAC TTAAAAACGT TGCTTTTATG
 721 AAAAGTTATA TTTTGGAAAA AGTAAAAGAA CACCAAGAAT CAATGGACAT GAACAACCCT
 781 CAGGACTTTA TTGATTGCTT CCTGATGAAA ATGGAGAAGG AAAAGCACAA CCAACCATCT
 841 GAATTTACTA TTGAAAGCTT GGAAAACACT GCAGTTGACT TGTTTGGAGC TGGGACAGAG
 901 ACGACAAGCA CAACCCTGAG ATATGCTCTC CTTCTCCTGC TGAAGCACCC AGAGGTCACA
 961 GCTAAAGTCC AGGAAGAGAT TGAACGTGTG ATTGGCAGAA ACCGGAGCCC CTGCATGCAA
1021 GACAGGAGCC ACATGCCCTA CACAGATGCT GTGGTGCACG AGGTCCAGAG ATACATTGAC
1081 CTTCTCCCCA CCAGCCTGCC CCATGCAGTG ACCTGTGACA TTAAATTCAG AAACTATCTC
1141 ATTCCCAAGG GCACAACCAT ATTAATTTCC CTGACTTCTG TGCTACATGA CAACAAAGAA
1201 TTTCCCAACC CAGAGATGTT TGACCCTCAT CACTTTCTGG ATGAAGGTGG CAATTTTAAG
1261 AAAAGTAAAT ACTTCATGCC TTTCTCAGCA GGAAAACGGA TTTGTGTGGG AGAAGCCCTG
1321 GCCGGCATGG AGCTGTTTTT ATTCCTGACC TCCATTTTAC AGAACTTTAA CCTGAAATCT
1381 CTGGTTGACC CAAAGAACCT TGACACCACT CCAGTTGTCA ATGGATTTGC CTCTGTGCCG
1441 CCCTTCTACC AGCTGTGCTT CATTCCTGTC TGAAGAAGAG CAGATGGCCT GGCTGCTGCT
1501 GTGCAGTCCC TGCAGCTCTC TTTCCTCTGG GGCATTATCC ATCTTTGCAC TATCTGTAAT
1561 GCCTTTTCTC ACCTGTCATC TCACATTTTC CCTTCCCTGA AGATCTAGTG AACATTCGAC
1621 CTCCATTACG GAGAGTTTCC TATGTTTCAC TGTGCAAATA TATCTGCTAT TCTCCATACT
1681 CTGTAACAGT TGCATTGACT GTCACATAAT GCTCATACTT ATCTAATGTA GAGTATTAAT
1741 ATGTTATTAT TAAATAGAGA AATATGATTT GTGTATTATA ATTCAAAGGC ATTTCTTTTC
1801 TGCATGATCT AAATAAAAAG CATTATTATT TGCTG
```

Figure 12A

```
  1 MDSLVVLVLC LSCLLLLSLW RQSSGRGKLP PGPTPLPVIG NILQIGIKDI SKSLTNLSKV
 61 YGPVFTLYFG LKPIVVLHGY EAVKEALIDL GEEFSGRGIF PLAERANRGF GIVFSNGKKW
121 KEIRRFSLMT LRNFGMGKRS IEDRVQEEAR CLVEELRKTK ASPCDPTFIL GCAPCNVICS
181 IIFHKRFDYK DQQFLNLMEK LNENIKILSS PWIQICNNFS PIIDYFPGTH NKLLKNVAFM
241 KSYILEKVKE HQESMDMNNP QDFIDCFLMK MEKEKHNQPS EFTIESLENT AVDLFGAGTE
301 TTSTTLRYAL LLLLKHPEVT AKVQEEIERV IGRNRSPCMQ DRSHMPYTDA VVHEVQRYID
361 LLPTSLPHAV TCDIKFRNYL IPKGTTILIS LTSVLHDNKE FPNPEMFDPH HFLDEGGNFK
421 KSKYFMPFSA GKRICVGEAL AGMELFLFLT SILQNFNLKS LVDPKNLDTT PVVNGFASVP
481 PFYQLCFIPV *RRADGLAAA VQSLQLSFLW GIIHLCTICN AFSHLSSHIF PSLKI**TFD
541 LHYGEFPMFH CANISAILHT L*QLH*LSHN AHTYLM*SIN MLLLNREI*F VYYNSKAFLF
601 CMI*IKSIII C
```

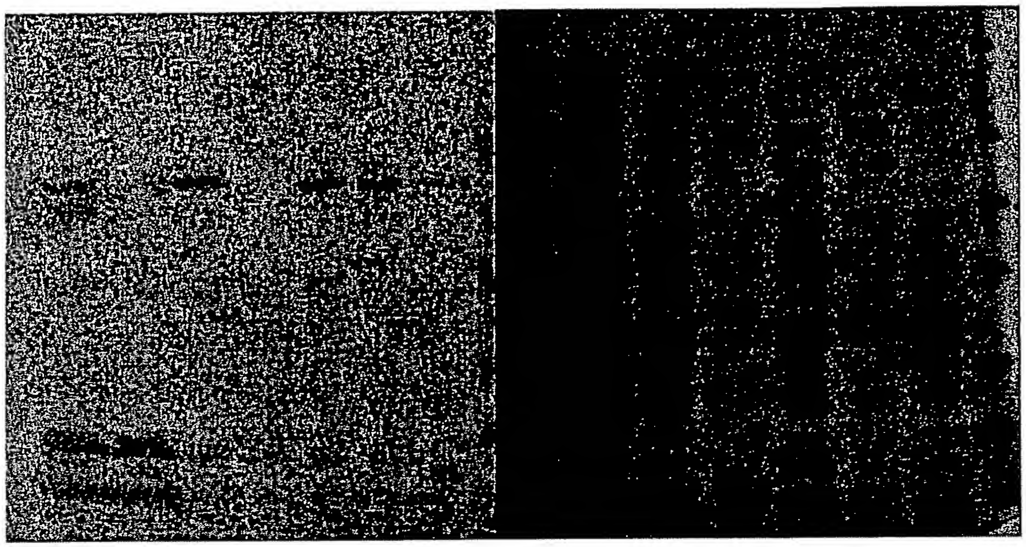
Figure 12B

1 ATGGGGCTAG AAGCACTGGT GCCCCTGGCC GTGATAGTGG CCATCTTCCT GCTCCTGGTG
61 GACCTGATGC ACCGGCGCCA ACGCTGGGCT GCACGCTACC CACCAGGCCC CCTGCCACTG
121 CCCGGGCTGG GCAACCTGCT GCATGTGGAC TTCCAGAACA CACCATACTG CTTCGACCAG
181 TTGCGGCGCC GCTTCGGGGA CGTGTTCAGC CTGCAGCTGG CCTGGACGCC GGTGGTCGTG
241 CTCAATGGGC TGGCGGCCGT GCGCGAGGCG CTGGTGACCC ACGGCGAGGA CACCGCCGAC
301 CGCCCGCCTG TGCCCATCAC CCAGATCCTG GGTTTCGGGC CGCGTTCCCA AGGGGTGTTC
361 CTGGCGCGCT ATGGGCCCGC GTGGCGCGAG CAGAGGCGCT TCTCCGTGTC CACCTTGCGC
421 AACTTGGGCC TGGGCAAGAA GTCGCTGGAG CAGTGGGTGA CCGAGGAGGC CGCCTGCCTT
481 TGTGCCGCCT TCGCCAACCA CTCCGGACGC CCCTTTCGCC CCAACGGTCT CTTGGACAAA
541 GCCGTGAGCA ACGTGATCGC CTCCCTCACC TGCGGGCGCC GCTTCGAGTA CGACGACCCT
601 CGCTTCCTCA GGCTGCTGGA CCTAGCTCAG GAGGGACTGA AGGAGGAGTC GGGCTTTCTG
661 CGCGAGGTGC TGAATGCTGT CCCCGTCCTC CTGCATATCC CAGCGCTGGC TGGCAAGGTC
721 CTACGCTTCC AAAAGGCTTT CCTGACCCAG CTGGATGAGC TGCTAACTGA GCACAGGATG
781 ACCTGGGACC CAGCCCAGCC CCCCCGAGAC CTGACTGAGG CCTTCCTGGC AGAGATGGAG
841 AAGGCCAAGG GGAACCCTGA GAGCAGCTTC AATGATGAGA ACCTGCGCAT AGTGGTGGCT
901 GACCTGTTCT CTGCCGGGAT GGTGACCACC TCGACCACGC TGGCCTGGGG CCTCCTGCTC
961 ATGATCCTAC ATCCGGATGT GCAGCGCCGT GTCCAACAGG AGATCGACGA CGTGATAGGG
1021 CAGGTGCGGC GACCAGAGAT GGGTGACCAG GCTCACATGC CCTACACCAC TGCCGTGATT
1081 CATGAGGTGC AGCGCTTTGG GGACATCGTC CCCCTGGGTA TGACCCATAT GACATCCCGT
1141 GACATCGAAG TACAGGGCTT CCGCATCCCT AAGGGAACGA CACTCATCAC CAACCTGTCA
1201 TCGGTGCTGA AGGATGAGGC CGTCTGGGAG AAGCCCTTCC GCTTCCACCC CGAACACTTC
1261 CTGGATGCCC AGGGCCACTT TGTGAAGCCG GAGGCCTTCC TGCCTTTCTC AGCAGGCCGC
1321 CGTGCATGCC TCGGGGAGCC CCTGGCCCGC ATGGAGCTCT TCCTCTTCTT CACCTCCCTG
1381 CTGCAGCACT TCAGCTTCTC GGTGCCCACT GGACAGCCCC GGCCCAGCCA CCATGGTGTC
1441 TTTGCTTTCC TGGTGAGCCC ATCCCCCTAT GAGCTTTGTG CTGTGCCCCG CTAG

Figure 13A

1 MGLEALVPLA VIVAIFLLLV DLMHRRQRWA ARYPPGPLPL PGLGNLLHVD FQNTPYCFDQ
61 LRRRFGDVFS LQLAWTPVVV LNGLAAVREA LVTHGEDTAD RPPVPITQIL GFGPRSQGVF
121 LARYGPAWRE QRRFSVSTLR NLGLGKKSLE QWVTEEAACL CAAFANHSGR PFRPNGLLDK
181 AVSNVIASLT CGRRFEYDDP RFLRLLDLAQ EGLKEESGFL REVLNAVPVL LHIPALAGKV
241 LRFQKAFLTQ LDELLTEHRM TWDPAQPPRD LTEAFLAEME KAKGNPESSF NDENLRIVVA
301 DLFSAGMVTT STTLAWGLLL MILHPDVQRR VQQEIDDVIG QVRRPEMGDQ AHMPYTTAVI
361 HEVQRFGDIV PLGMTHMTSR DIEVQGFRIP KGTTLITNLS SVLKDEAVWE KPFRFHPEHF
421 LDAQGHFVKP EAFLPFSAGR RACLGEPLAR MELFLFFTSL LQHFSFSVPT GQPRPSHHGV
481 FAFLVSPSPY ELCAVPR*

Figure 13B



Lane1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8

Figure 14

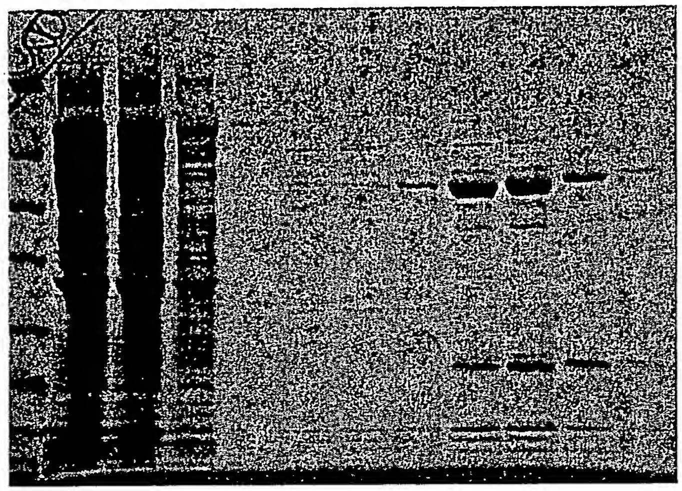


Figure 15

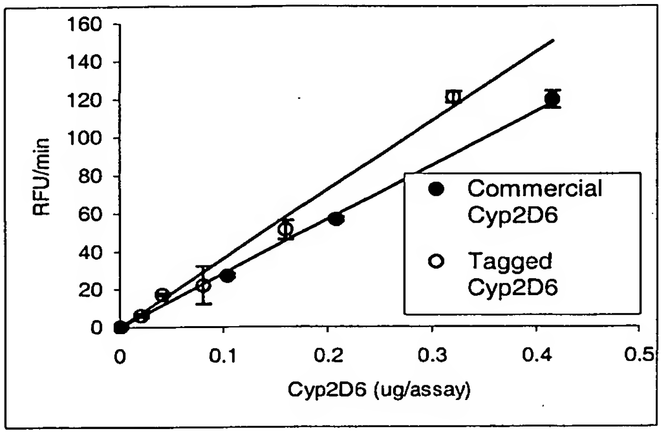

160
140
120
100
80
60
40
20
0
RFU/min
0
0.1
0.2
0.3
0.4
0.5
Cyp2D6 (ug/assay)
Commercial Cyp2D6
Tagged Cyp2D6

Figure 16

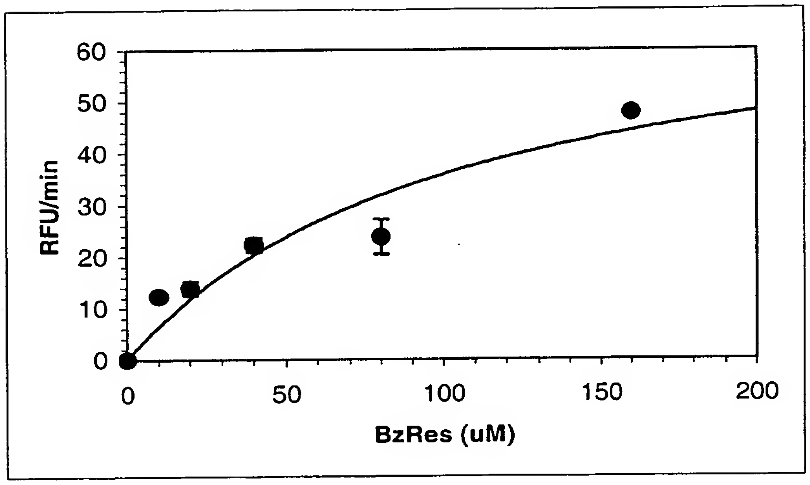


Figure 17

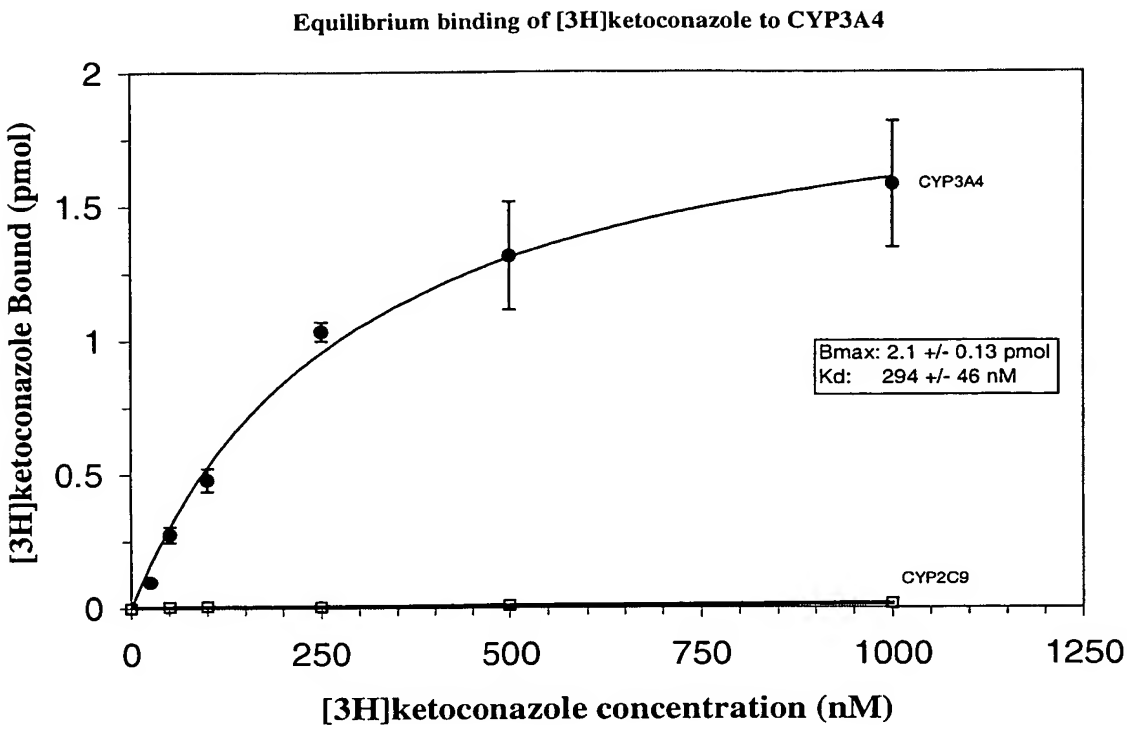
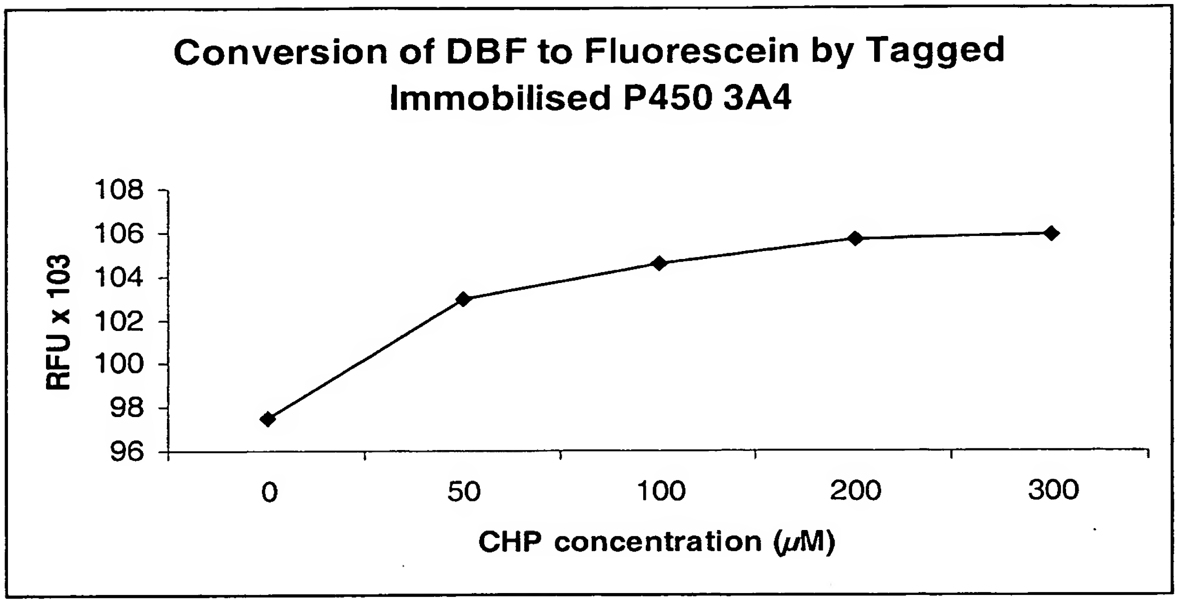

Equilibrium binding of [3H]ketoconazole to CYP3A4
[3H]ketoconazole Bound (pmol)
2
1.5
1
0.5
0
0
250
500
750
1000
1250
[3H]ketoconazole concentration (nM)
CYP3A4
CYP2C9
Bmax: 2.1 +/- 0.13 pmol
Kd: 294 +/- 46 nM

Figure 18



Conversion of DBF to Fluorescein by Tagged Immobilised P450 3A4
RFU x 103
108
106
104
102
100
98
96
0
50
100
200
300
CHP concentration (μM)

Figure 19

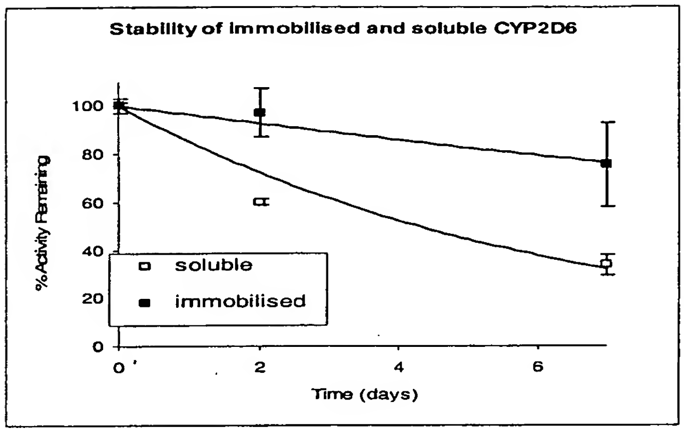

Stability of immobilised and soluble CYP2D6
100
80
60
40
20
0
% Activity Remaining
0
2
4
6
Time (days)
soluble
immobilised

Figure 20

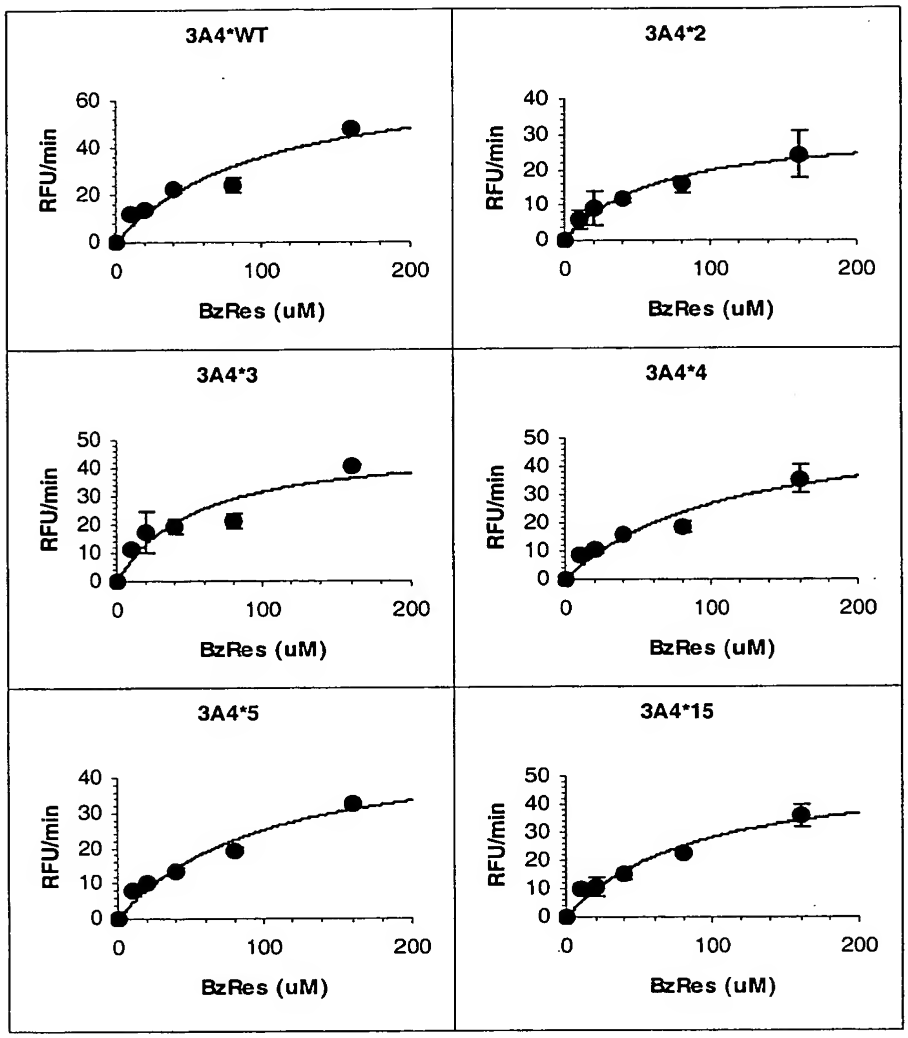
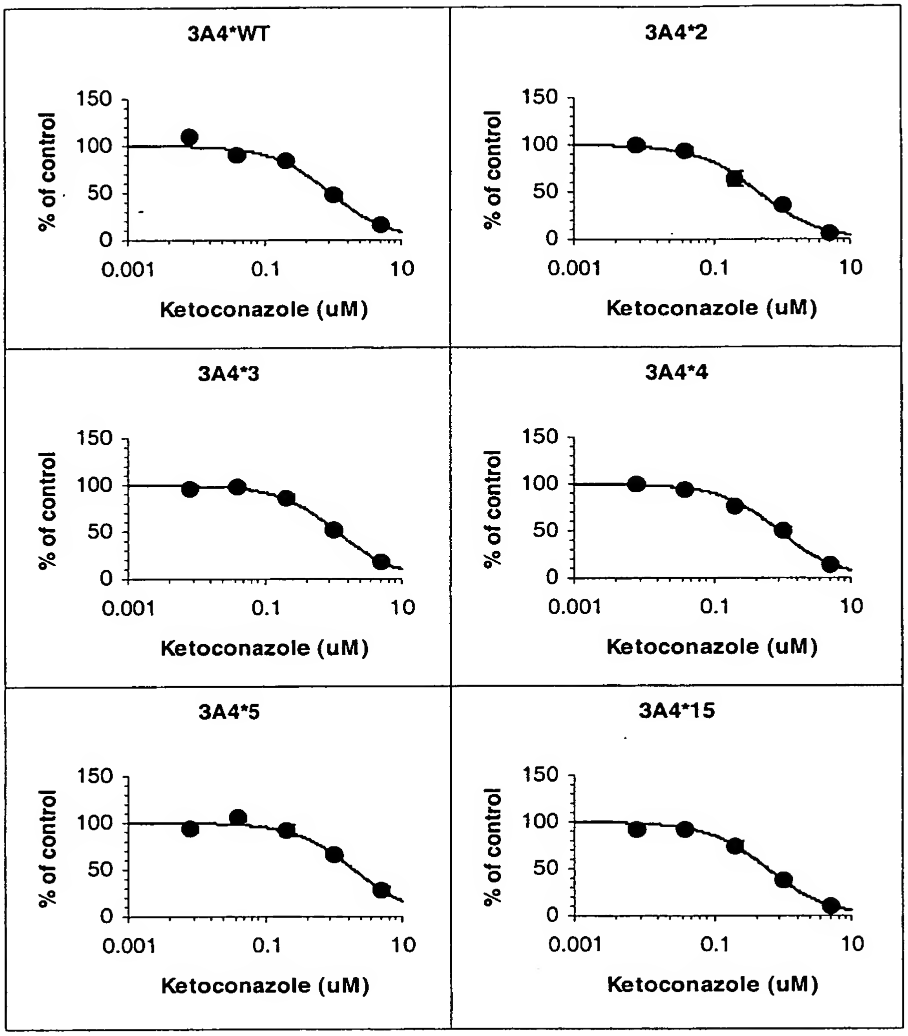

3A4*WT
RFU/min
BzRes (uM)
3A4*2
RFU/min
BzRes (uM)
3A4*3
RFU/min
BzRes (uM)
3A4*4
RFU/min
BzRes (uM)
3A4*5
RFU/min
BzRes (uM)
3A4*15
RFU/min
BzRes (uM)

Figure 21



3A4*WT
3A4*2
3A4*3
3A4*4
3A4*5
3A4*15
% of control
150
100
50
0
0.001
0.1
10
Ketoconazole (uM)

Figure 22

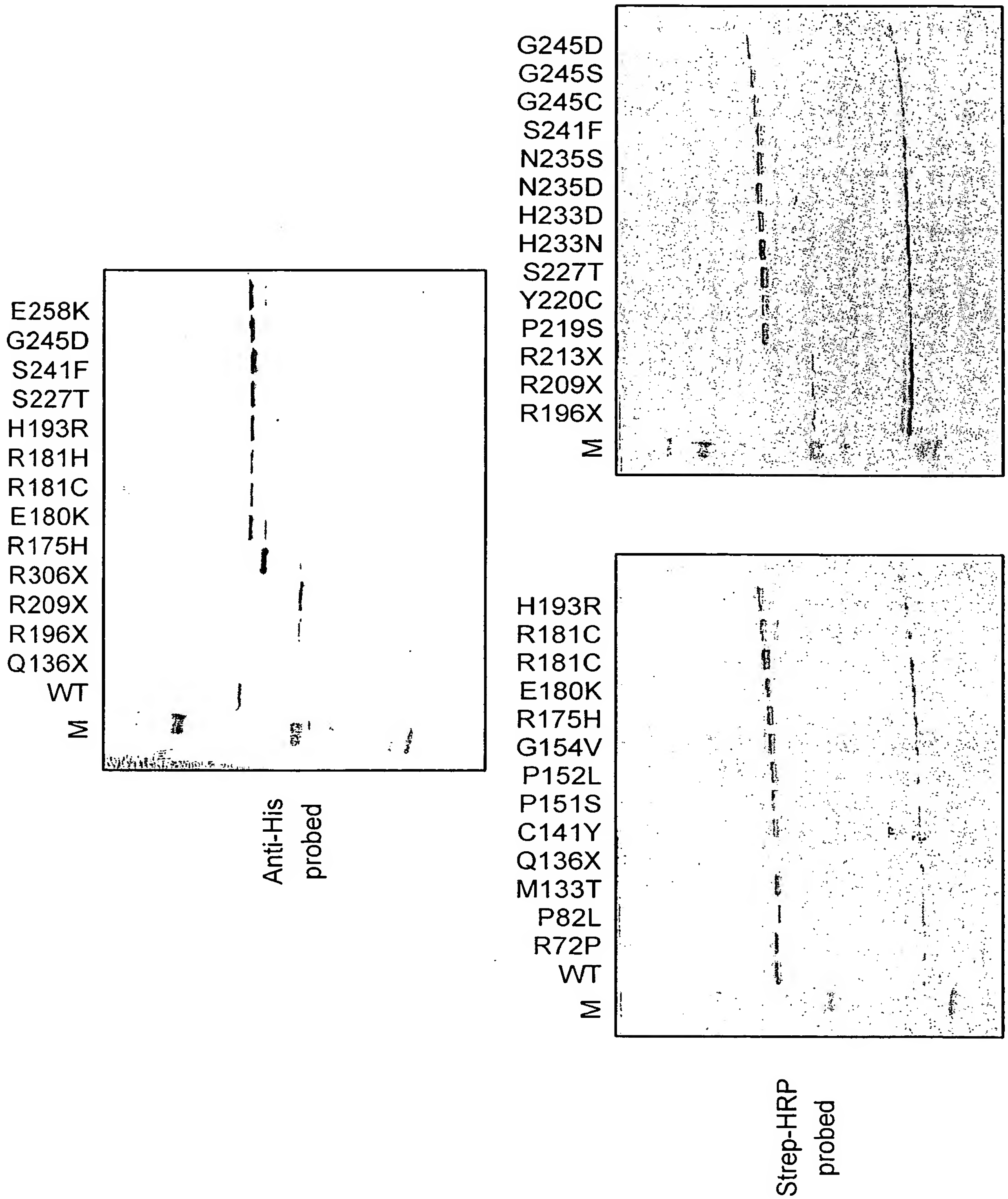

E258K
G245D
S241F
S227T
H193R
R181H
R181C
E180K
R175H
R306X
R209X
R196X
Q136X
WT
M
Anti-His probed
G245D
G245S
G245C
S241F
N235S
N235D
H233D
H233N
S227T
Y220C
P219S
R213X
R209X
R196X
M
H193R
R181C
R181C
E180K
R175H
G154V
P152L
P151S
C141Y
Q136X
M133T
P82L
R72P
WT
M
Strep-HRP probed

FIG. 1

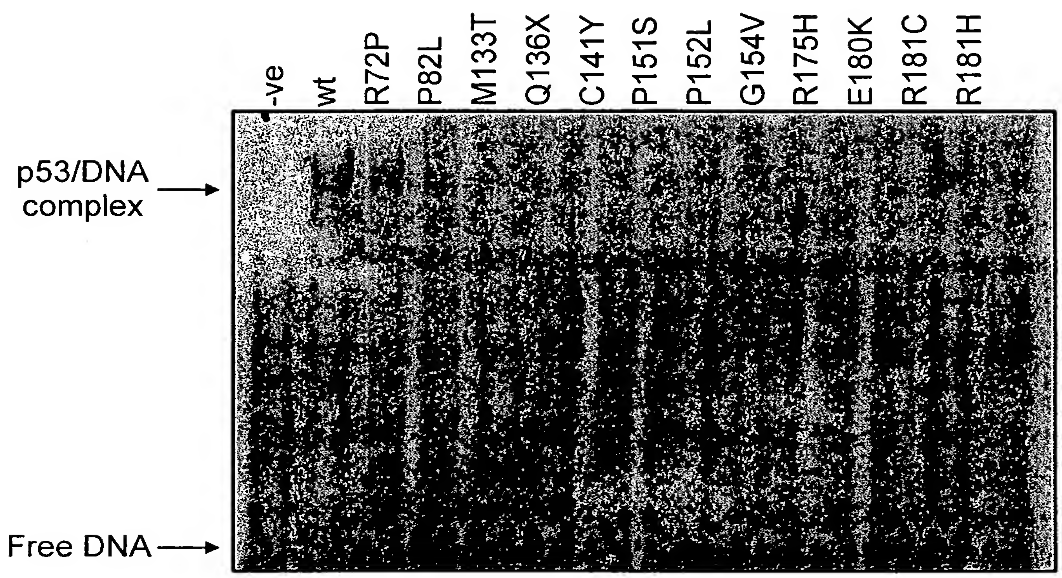

-ve
wt
R72P
P82L
M133T
Q136X
C141Y
P151S
P152L
G154V
R175H
E180K
R181C
R181H
p53/DNA complex
Free DNA

FIG. 2

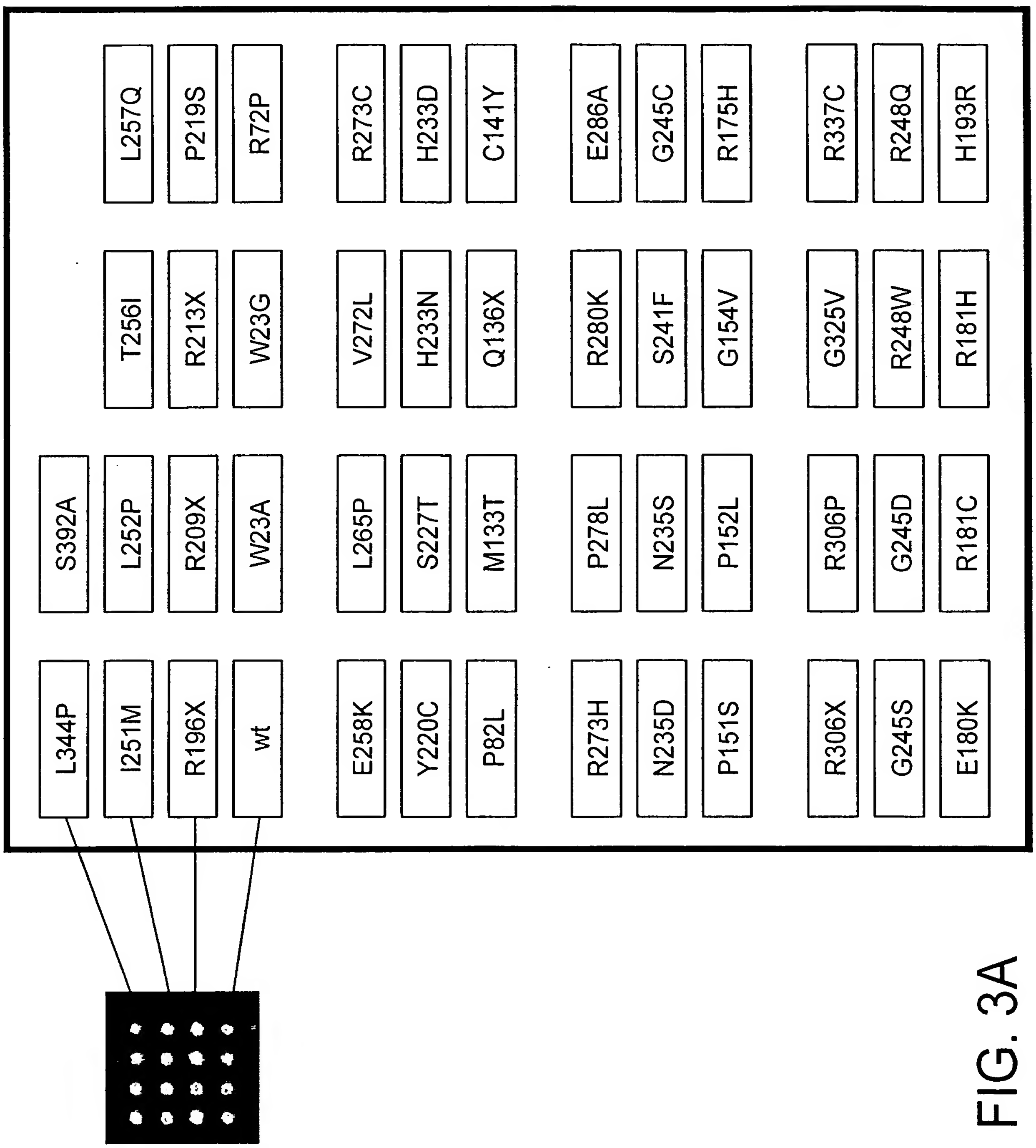

L257Q
P219S
R72P
R273C
H233D
C141Y
E286A
G245C
R175H
R337C
R248Q
H193R
T256I
R213X
W23G
V272L
H233N
Q136X
R280K
S241F
G154V
G325V
R248W
R181H
S392A
L252P
R209X
W23A
L265P
S227T
M133T
P278L
N235S
P152L
R306P
G245D
R181C
L344P
I251M
R196X
wt
E258K
Y220C
P82L
R273H
N235D
P151S
R306X
G245S
E180K

FIG. 3A

PCT/GB 2002 / 0 0 5 4 9 9
2 6 FEB 2003

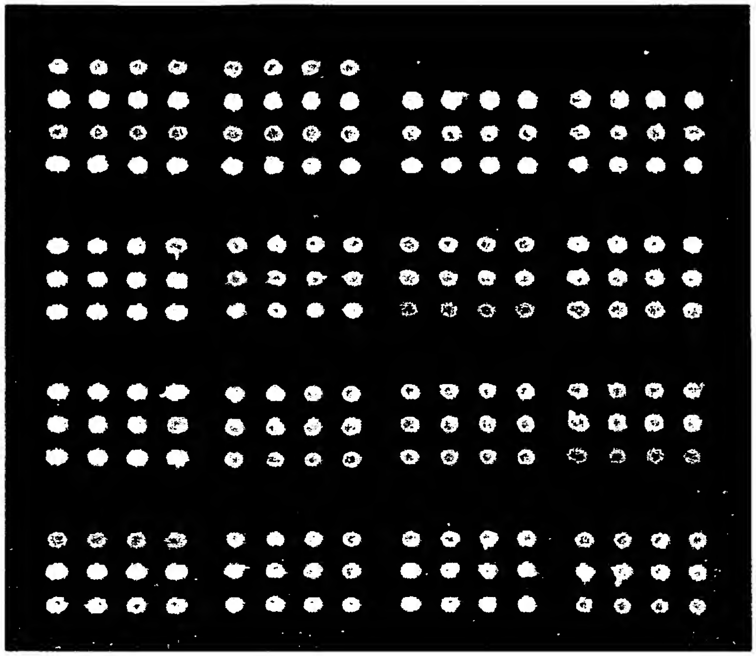


FIG. 3B

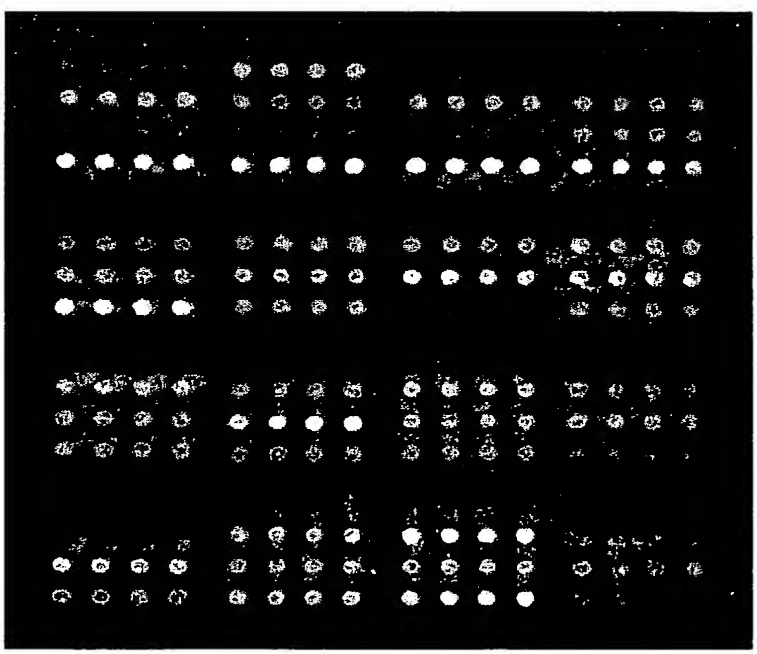


FIG. 3C

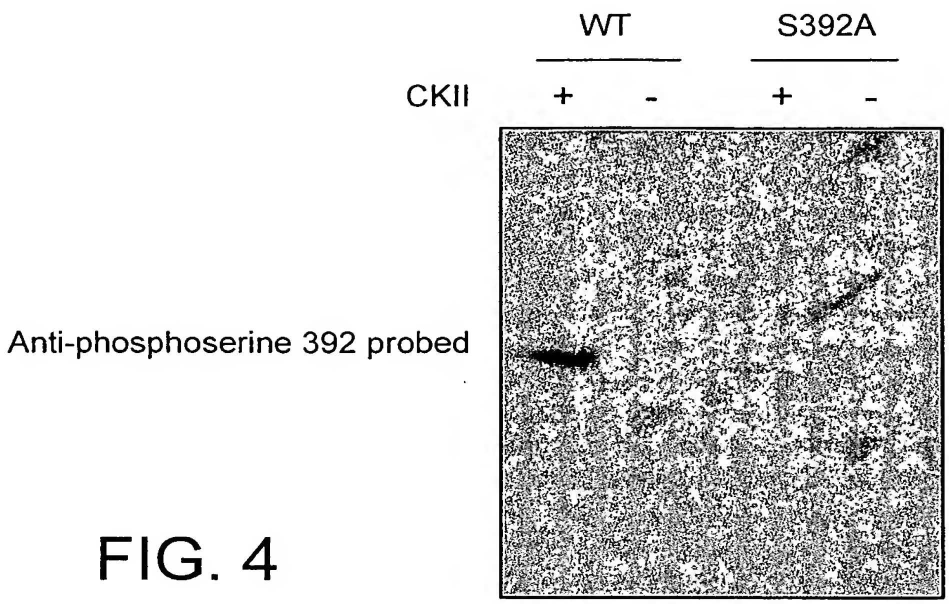


FIG. 4

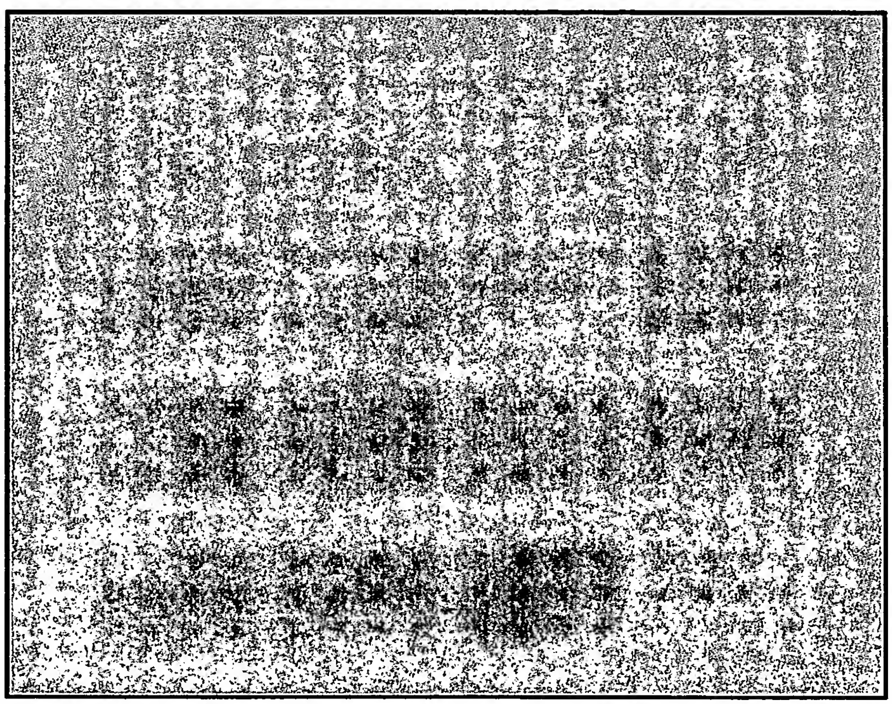


FIG. 5

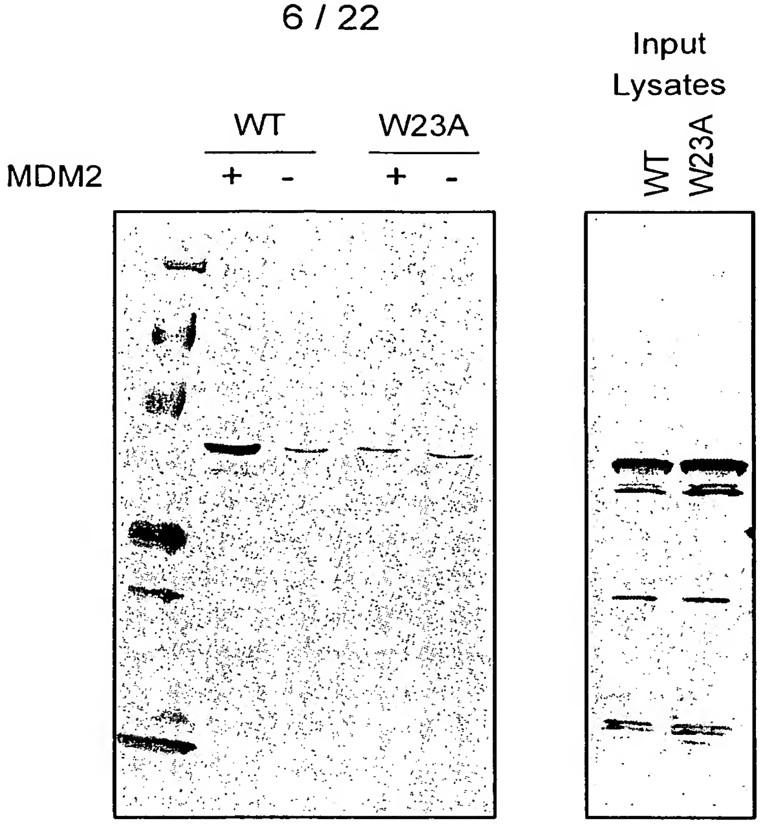


FIG. 6

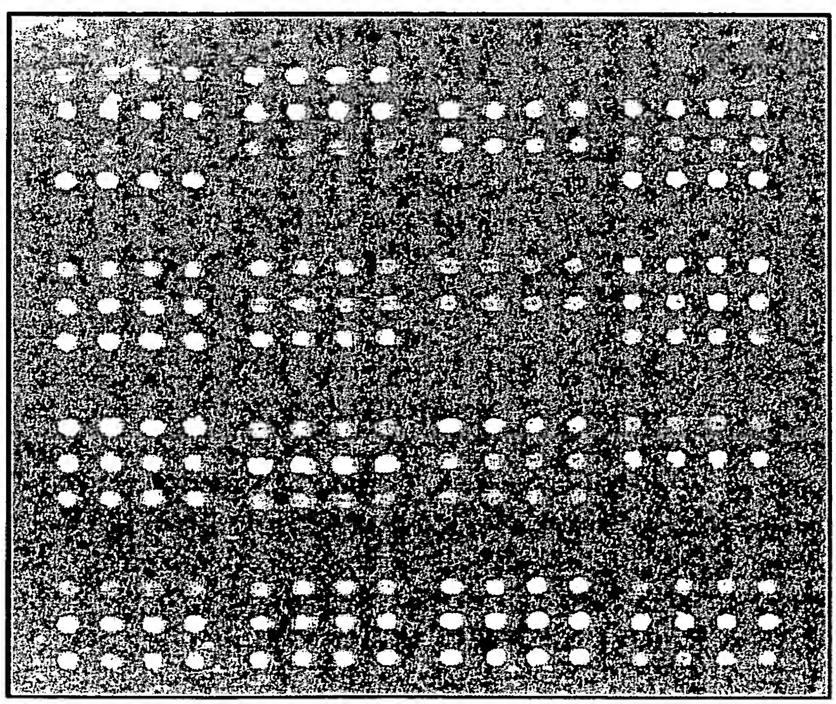


FIG. 7

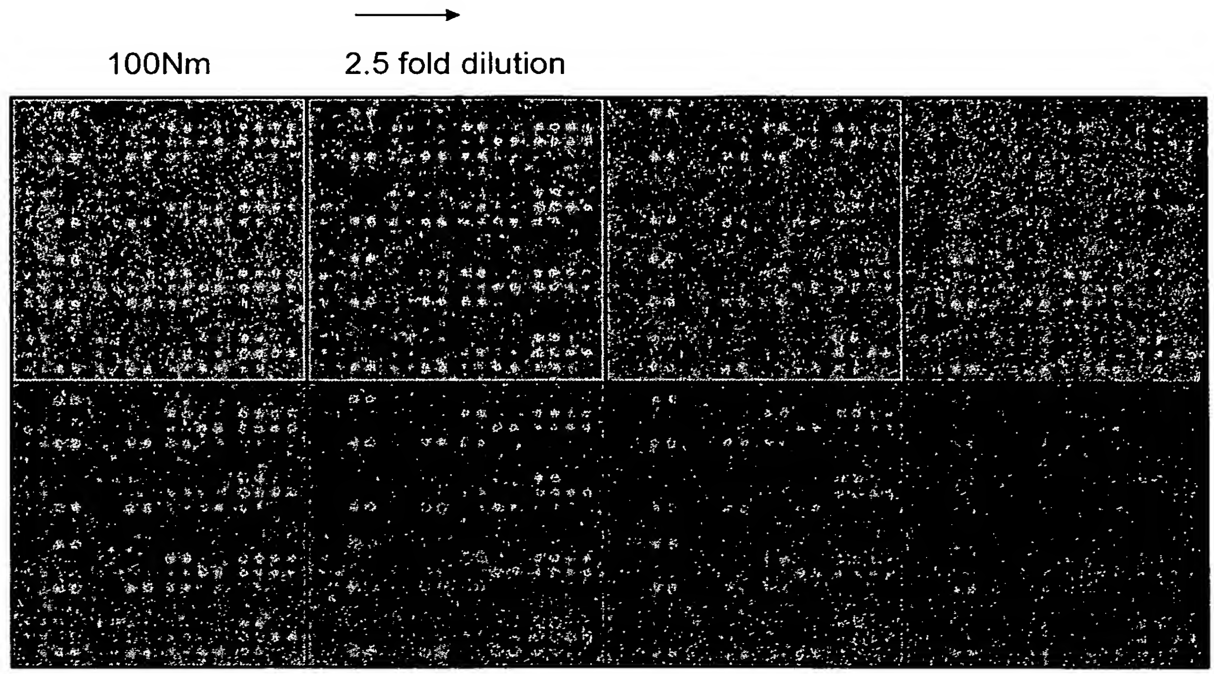

100Nm
2.5 fold dilution

FIG. 8A

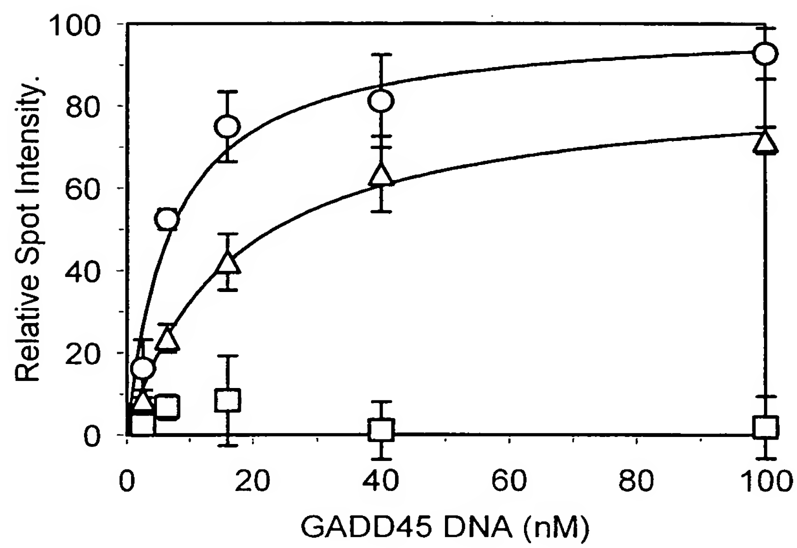

Relative Spot Intensity.
0
20
40
60
80
100
GADD45 DNA (nM)

FIG. 8B

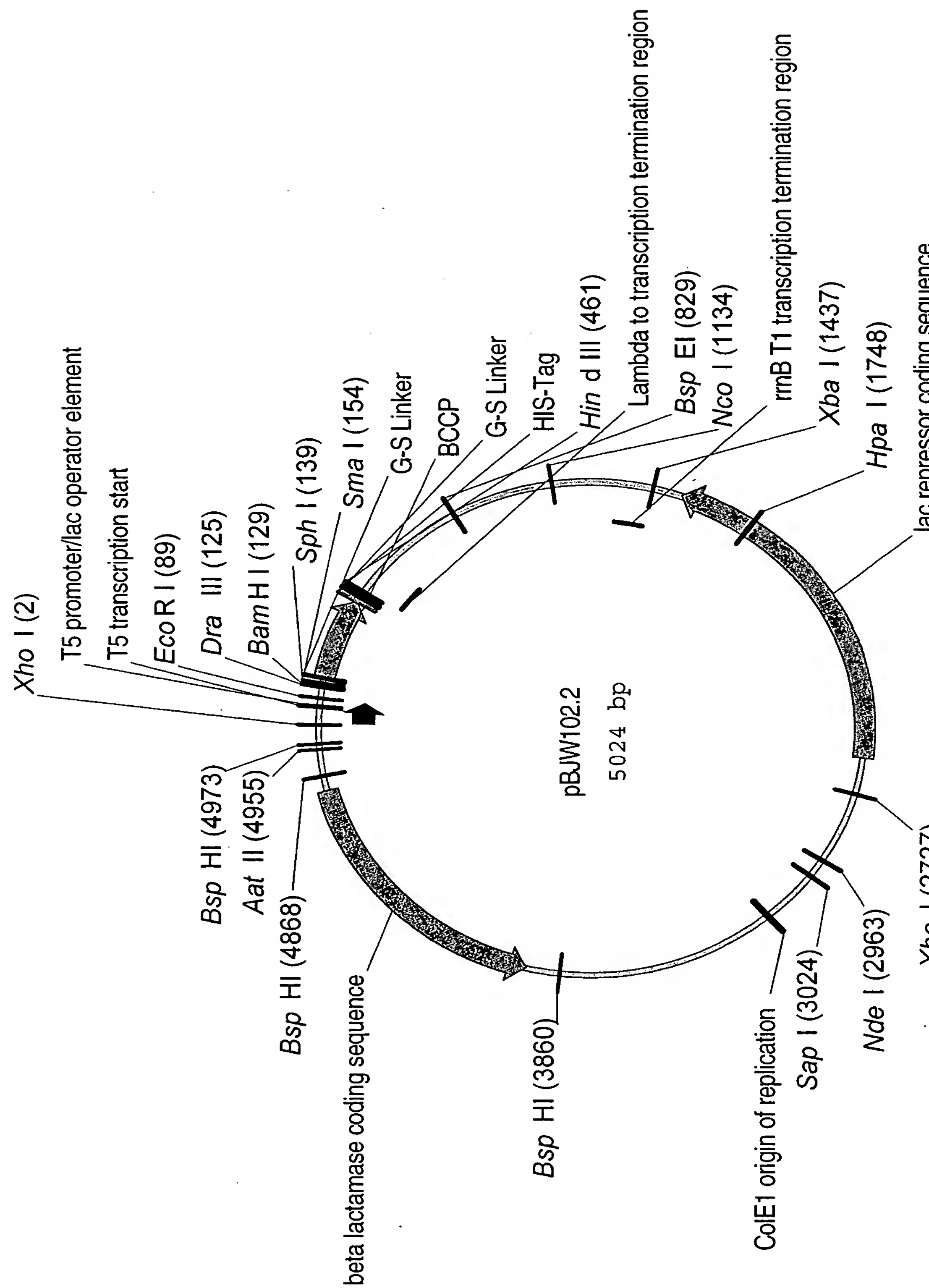

Xho I (2)
T5 promoter/lac operator element
T5 transcription start
Eco R I (89)
Dra III (125)
Bam H I (129)
Sph I (139)
Sma I (154)
G-S Linker
BCCP
G-S Linker
HIS-Tag
Hin d III (461)
Lambda to transcription termination region
Bsp EI (829)
Nco I (1134)
rrnB T1 transcription termination region
Xba I (1437)
Hpa I (1748)
lac repressor coding sequence
Xba I (2727)
Nde I (2963)
Sap I (3024)
ColE1 origin of replication
Bsp HI (3860)
beta lactamase coding sequence
Bsp HI (4868)
Aat II (4955)
Bsp HI (4973)
pBJW102.2
5024 bp

FIG. 9A

```
   1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA
  61 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGGCA
 121 CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGCGGTTC TGGCGCAGCA
 181 GCGGAAATCA GTGGTCACAT CGTACGTTCC CCGATGGTTG GTACTTTCTA CCGCACCCCA
 241 AGCCCGGACG CAAAAGCGTT CATCGAAGTG GGTCAGAAAG TCAACGTGGG CGATACCCTG
 301 TGCATCGTTG AAGCCATGAA AATGATGAAC CAGATCGAAG CGGACAAATC CGGTACCGTG
 361 AAAGCAATTC TGGTCGAAAG TGGACAACCG GTAGAATTTG ACGAGCCGCT GGTCGTCATC
 421 GAGGGTGGCA GCGGTTCTGG CCACCATCAC CATCACCATA AGCTTAATTA GCTGAGCTTG
 481 GACTCCTGTT GATAGATCCA GTAATGACCT CAGAACTCCA TCTGGATTTG TTCAGAACGC
 541 TCGGTTGCCG CCGGGCGTTT TTTATTGGTG AGAATCCAAG CTAGCTTGGC GAGATTTTCA
 601 GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA ATCACTGGAT ATACCACCGT TGATATATCC
 661 CAATGGCATC GTAAAGAACA TTTTGAGGCA TTTCAGTCAG TTGCTCAATG TACCTATAAC
 721 CAGACCGTTC AGCTGGATAT TACGGCCTTT TTAAAGACCG TAAAGAAAAA TAAGCACAAG
 781 TTTTATCCGG CCTTTATTCA CATTCTTGCC CGCCTGATGA ATGCTCATCC GGAATTTCGT
 841 ATGGCAATGA AAGACGGTGA GCTGGTGATA TGGGATAGTG TTCACCCTTG TTACACCGTT
 901 TTCCATGAGC AAACTGAAAC GTTTTCATCG CTCTGGAGTG AATACCACGA CGATTTCCGG
 961 CAGTTTCTAC ACATATATTC GCAAGATGTG GCGTGTTACG GTGAAAACCT GGCCTATTTC
1021 CCTAAAGGGT TTATTGAGAA TATGTTTTTC GTCTCAGCCA ATCCCTGGGT GAGTTTCACC
1081 AGTTTTGATT TAAACGTGGC CAATATGGAC AACTTCTTCG CCCCCGTTTT CACCATGGGC
1141 AAATATTATA CGCAAGGCGA CAAGGTGCTG ATGCCGCTGG CGATTCAGGT TCATCATGCC
1201 GTTTGTGATG GCTTCCATGT CGGCAGAATG CTTAATGAAT TACAACAGTA CTGCGATGAG
1261 TGGCAGGGCG GGGCGTAATT TTTTTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGGGTA
1321 ATGACTCTCT AGCTTGAGGC ATCAAATAAA ACGAAAGGCT CAGTCGAAAG ACTGGGCCTT
1381 TCGTTTTATC TGTTGTTTGT CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCCTCTAG
1441 ATTACGTGCA GTCGATGATA AGCTGTCAAA CATGAGAATT GTGCCTAATG AGTGAGCTAA
1501 CTTACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT CGGGAAACCT GTCGTGCCAG
1561 CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCCAGGGT
1621 GGTTTTTCTT TTCACCAGTG AGACGGGCAA CAGCTGATTG CCCTTCACCG CCTGGCCCTG
1681 AGAGAGTTGC AGCAAGCGGT CCACGCTGGT TTGCCCCAGC AGGCGAAAAT CCTGTTTGAT
1741 GGTGGTTAAC GGCGGGATAT AACATGAGCT GTCTTCGGTA TCGTCGTATC CCACTACCGA
1801 GATATCCGCA CCAACGCGCA GCCCGGACTC GGTAATGGCG CGCATTGCGC CCAGCGCCAT
1861 CTGATCGTTG GCAACCAGCA TCGCAGTGGG AACGATGCCC TCATTCAGCA TTTGCATGGT
1921 TTGTTGAAAA CCGGACATGG CACTCCAGTC GCCTTCCCGT TCCGCTATCG GCTGAATTTG
1981 ATTGCGAGTG AGATATTTAT GCCAGCCAGC CAGACGCAGA CGCGCCGAGA CAGAACTTAA
2041 TGGGCCCGCT AACAGCGCGA TTTGCTGGTG ACCCAATGCG ACCAGATGCT CCACGCCCAG
2101 TCGCGTACCG TCTTCATGGG AGAAAATAAT ACTGTTGATG GGTGTCTGGT CAGAGACATC
2161 AAGAAATAAC GCCGGAACAT TAGTGCAGGC AGCTTCCACA GCAATGGCAT CCTGGTCATC
2221 CAGCGGATAG TTAATGATCA GCCCACTGAC GCGTTGCGCG AGAAGATTGT GCACCGCCGC
2281 TTTACAGGCT TCGACGCCGC TTCGTTCTAC CATCGACACC ACCACGCTGG CACCCAGTTG
2341 ATCGGCGCGA GATTTAATCG CCGCGACAAT TTGCGACGGC GCGTGCAGGG CCAGACTGGA
2401 GGTGGCAACG CCAATCAGCA ACGACTGTTT GCCCGCCAGT TGTTGTGCCA CGCGGTTGGG
2461 AATGTAATTC AGCTCCGCCA TCGCCGCTTC CACTTTTTCC CGCGTTTTCG CAGAAACGTG
2521 GCTGGCCTGG TTCACCACGC GGGAAACGGT CTGATAAGAG ACACCGGCAT ACTCTGCGAC
2581 ATCGTATAAC GTTACTGGTT TCACATTCAC CACCCTGAAT TGACTCTCTT CCGGGCGCTA
2641 TCATGCCATA CCGCGAAAGG TTTTGCACCA TTCGATGGTG TCGGAATTTC GGGCAGCGTT
2701 GGGTCCTGGC CACGGGTGCG CATGATCTAG AGCTGCCTCG CGCGTTTCGG TGATGACGGT
2761 GAAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG CTTGTCTGTA AGCGGATGCC
2821 GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA GCGGGTGTTG GCGGGTGTCG GGGCGCAGCC
2881 ATGACCCAGT CACGTAGCGA TAGCGGAGTG TATACTGGCT TAACTATGCG GCATCAGAGC
2941 AGATTGTACT GAGAGTGCAC CATATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA
3001 AATACCGCAT CAGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC
3061 GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG
3121 GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
3181 AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
3241 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC
3301 CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG
3361 CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT
3421 CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC
3481 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
3541 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
3601 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
```

FIG. 9B

```
3661 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
3721 CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG
3781 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
3841 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA
3901 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
3961 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG
4021 TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
4081 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
4141 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
4201 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
4261 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
4321 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
4381 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
4441 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
4501 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
4561 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
4621 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
4681 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
4741 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
4801 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
4861 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC
4921 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
4981 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT TCAC
```

FIG. 9B CONT'D

```
          Dra III             Sph I                  Sma I
115   ATGGCA CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGC
      TACCGT GAATCACCCT AGGCGTACGC TCGAGCCATG GGGCCCCCAC CGTCG
```

FIG. 9C

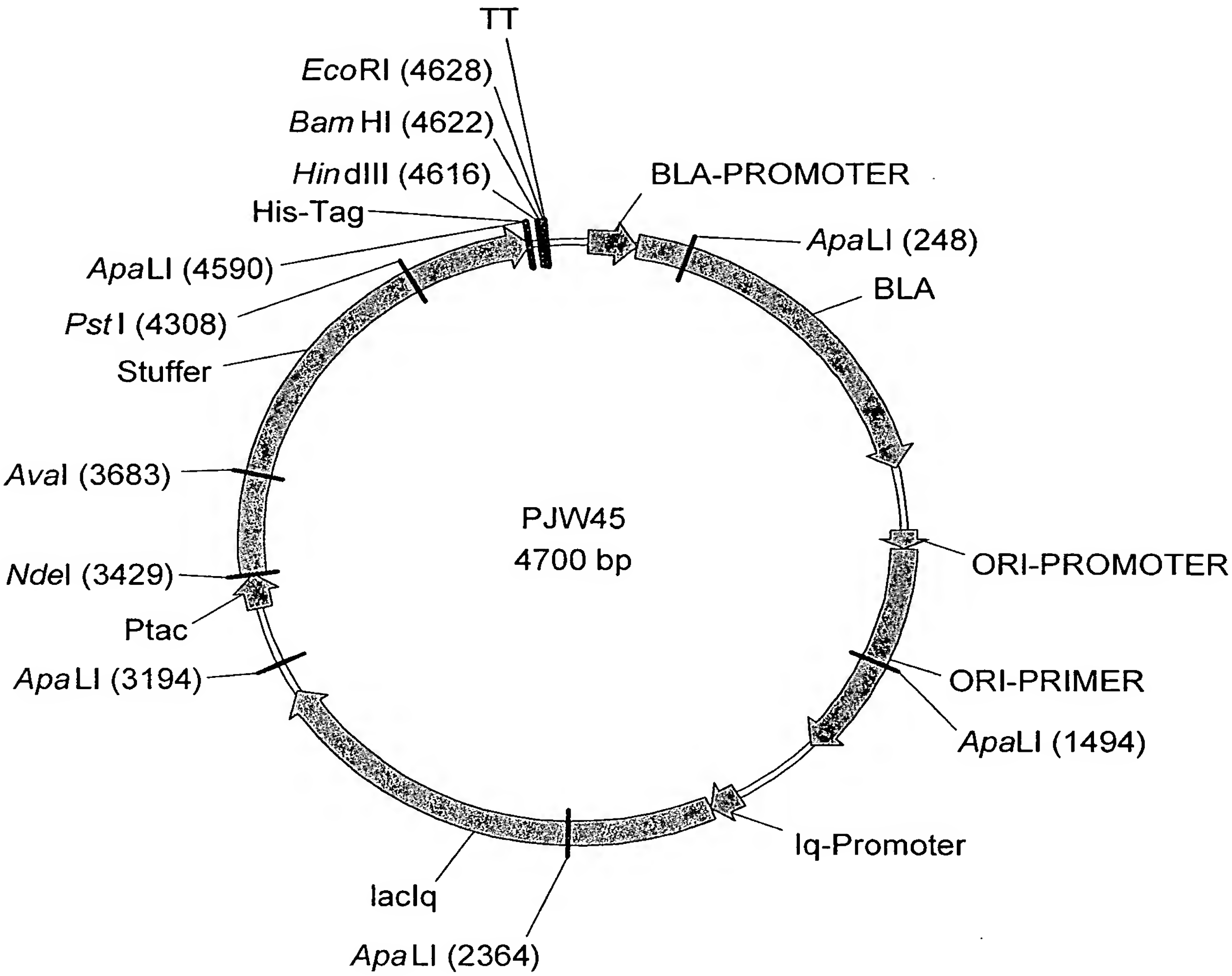

TT
EcoRI (4628)
Bam HI (4622)
HindIII (4616)
His-Tag
ApaLI (4590)
PstI (4308)
Stuffer
AvaI (3683)
NdeI (3429)
Ptac
ApaLI (3194)
BLA-PROMOTER
ApaLI (248)
BLA
PJW45
4700 bp
ORI-PROMOTER
ORI-PRIMER
ApaLI (1494)
Iq-Promoter
lacIq
ApaLI (2364)

FIG. 10A

1 CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
61 ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA
121 AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT
181 TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
241 AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
301 GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG
361 CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
421 AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG
481 TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
541 TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG
601 TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG
661 ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
721 TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC
781 CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
841 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG
901 TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
961 AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC
1021 TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG
1081 ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
1141 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC
1201 AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
1261 TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT
1321 AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC
1381 TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
1441 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC
1501 AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCATTGAG
1561 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG
1621 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
1681 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
1741 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT
1801 TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
1861 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG
1921 AGGAAGCCCA GGACCCAACG CTGCCCGAAA TTCCGACACC ATCGAATGGT GCAAAACCTT
1981 TCGCGGTATG GCATGATAGC GCCCGGAAGA GAGTCAATTC AGGGTGGTGA ATGTGAAACC
2041 AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCG TTTCCCGCGT
2101 GGTGAACCAG GCCAGCCACG TTTCTGCGAA AACGCGGGAA AAAGTGGAAG CGGCGATGGC
2161 GGAGCTGAAT TACATTCCCA ACCGCGTGGC ACAACAACTG GCGGGCAAAC AGTCGTTGCT
2221 GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACGCGCCG TCGCAAATTG TCGCGGCGAT
2281 TAAATCTCGC GCCGATCAAC TGGGTGCCAG CGTGGTGGTG TCGATGGTAG AACGAAGCGG
2341 CGTCGAAGCC TGTAAAGCGG CGGTGCACAA TCTTCTCGCG CAACGCGTCA GTGGGCTGAT
2401 CATTAACTAT CCGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT GCACTAATGT
2461 TCCGGCGTTA TTTCTTGATG TCTCTGACCA GACACCCATC AACAGTATTA TTTTCTCCCA

FIG. 10B

```
2521 TGAAGACGGT ACGCGACTGG GCGTGGAGCA TCTGGTCGCA TTGGGTCACC AGCAAATCGC
2581 GCTGTTAGCG GGCCCATTAA GTTCTGTCTC GGCGCGTCTG CGTCTGGCTG GCTGGCATAA
2641 ATATCTCACT CGCAATCAAA TTCAGCCGAT AGCGGAACGG GAAGGCGACT GGAGTGCCAT
2701 GTCCGGTTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCCA CTGCGATGCT
2761 GGTTGCCAAC GATCAGATGG CGCTGGGCGC AATGCGCGCC ATTACCGAGT CCGGGCTGCG
2821 CGTTGGTGCG GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT CATGTTATAT
2881 CCCGCCGTTA ACCACCATCA AACAGGATTT TCGCCTGCTG GGGCAAACCA GCGTGGACCG
2941 CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC CCGTCTCACT
3001 GGTGAAAAGA AAAACCACCC TGGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC
3061 CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA
3121 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACAATT CTCATGTTTG ACAGCTTATC
3181 ATCGACTGCA CGGTGCACCA ATGCTTCTGG CGTCAGGCAG CCATCGGAAG CTGTGGTATG
3241 GCTGTGCAGG TCGTAAATCA CTGCATAATT CGTGTCGCTC AAGGCGCACT CCCGTTCTGG
3301 ATAATGTTTT TTGCGCCGAC ATCATAACGG TTCTGGCAAA TATTCTGAAA TGAGCTGTTG
3361 ACAATTAATC ATCGGCTCGT ATAATGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG
3421 GAAACACATA TGAACGACTT TCATCGCGAT ACGTGGGCGG AAGTGGATTT GGACGCCATT
3481 TACGACAATG TGGCGAATTT GCGCCGTTTG CTGCCGGACG ACACGCACAT TATGGCGGTC
3541 GTGAAGGCGA ACGCCTATGG ACATGGGGAT GTGCAGGTGG CAAGGACAGC GCTCGAAGCG
3601 GGGGCCTCCC GCCTGGCGGT TGCCTTTTTG GATGAGGCGC TCGCTTTAAG GGAAAAAGGA
3661 ATCGAAGCGC CGATTCTAGT TCTCGGGGCT TCCCGTCCAG CTGATGCGGC GCTGGCCGCC
3721 CAGCAGCGCA TTGCCCTGAC CGTGTTCCGC TCCGACTGGT TGGAAGAAGC GTCCGCCCTT
3781 TACAGCGGCC CTATTCCTAT TCATTTCCAT TTGAAAATGG ACACCGGCAT GGGACGGCTT
3841 GGAGTGAAAG ACGAGGAGGA GACGAAACGA ATCGCAGCGC TGATTGAGCG CCATCCGCAT
3901 TTTGTGCTTG AAGGGGCGTA CACGCATTTT GCGACTGCGG ATGAGGTGAA CACCGATTAT
3961 TTTTCCTATC AGTATACCCG TTTTTTGCAC ATGCTCGAAT GGCTGCCGTC GCGCCCGCCG
4021 CTCGTCCATT GCGCCAACAG CGCAGCGTCG CTCCGTTTCC CTGACCGGAC GTTCAATATG
4081 GTCCGCTTCG GCATTGCCAT GTATGGGCTT GCCCCGTCGC CCGGCATCAA GCCGCTGCTG
4141 CCGTATCCAT TAAAAGAAGC ATTTTCGCTC CATAGCCGCC TCGTACACGT CAAAAAACTG
4201 CAACCAGGCG AAAAGGTGAG CTATGGTGCG ACGTACACTG CGCAGACGGA GGAGTGGATC
4261 GGGACGATTC CGATCGGCTA TGCGGACGGC TGGCTCCGCC GCCTGCAGCA CTTTCATGTC
4321 CTTGTTGACG GACAAAAGGC GCCGATTGTC GGCCGCATTT GCATGGACCA GTGCATGATC
4381 CGCCTGCCTG GGCCGCTGCC GGTCGGCACG AAGGTGACAC TGATTGGTCG CCAGGGGGAC
4441 GAGGTAATTT CCATTGATGA TGTCGCTCGC CATTTGGAAA CGATCAACTA CGAAGTGCCT
4501 TGCACGATCA GCTATCGAGT GCCCCGTATT TTTTTCCGCC ATAAGCGTAT AATGGAAGTG
4561 AGAAACGCCA TTGGCCGCGG GGAAAGCAGT GCACATCACC ATCACCATCA CTAAAAGCTT
4621 GGATCCGAAT TCAGCCCGCC TAATGAGCGG GCTTTTTTTT GAACAAAATT AGCTTGGCTG
4681 TTTTGGCGGA TGAGAGAAGA
```

FIG. 10B CONT'D

```
   1 ATGGCTCTCA TCCCAGACTT GGCCATGGAA ACCTGGCTTC TCCTGGCTGT CAGCCTGGTG
  61 CTCCTCTATC TATATGGAAC CCATTCACAT GGACTTTTTA AGAAGCTTGG AATTCCAGGG
 121 CCCACACCTC TGCCTTTTTT GGGAAATATT TTGTCCTACC ATAAGGGCTT TTGTATGTTT
 181 GACATGGAAT GTCATAAAAA GTATGGAAAA GTGTGGGGCT TTTATGATGG TCAACAGCCT
 241 GTGCTGGCTA TCACAGATCC TGACATGATC AAAACAGTGC TAGTGAAAGA ATGTTATTCT
 301 GTCTTCACAA ACCGGAGGCC TTTTGGTCCA GTGGGATTTA TGAAAAGTGC CATCTCTATA
 361 GCTGAGGATG AAGAATGGAA GAGATTACGA TCATTGCTGT CTCCAACCTT CACCAGTGGA
 421 AAACTCAAGG AGATGGTCCC TATCATTGCC CAGTATGGAG ATGTGTTGGT GAGAAATCTG
 481 AGGCGGGAAG CAGAGACAGG CAAGCCTGTC ACCTTGAAAG ACGTCTTTGG GGCCTACAGC
 541 ATGGATGTGA TCACTAGCAC ATCATTTGGA GTGAACATCG ACTCTCTCAA CAATCCACAA
 601 GACCCCTTTG TGGAAAACAC CAAGAAGCTT TTAAGATTTG ATTTTTTGGA TCCATTCTTT
 661 CTCTCAATAA CAGTCTTTCC ATTCCTCATC CCAATTCTTG AAGTATTAAA TATCTGTGTG
 721 TTTCCAAGAG AAGTTACAAA TTTTTTAAGA AAATCTGTAA AAAGGATGAA AGAAAGTCGC
 781 CTCGAAGATA CACAAAAGCA CCGAGTGGAT TTCCTTCAGC TGATGATTGA CTCTCAGAAT
 841 TCAAAAGAAA CTGAGTCCCA CAAAGCTCTG TCCGATCTGG AGCTCGTGGC CCAATCAATT
 901 ATCTTTATTT TTGCTGGCTA TGAAACCACG AGCAGTGTTC TCTCCTTCAT TATGTATGAA
 961 CTGGCCACTC ACCCTGATGT CCAGCAGAAA CTGCAGGAGG AAATTGATGC AGTTTTACCC
1021 AATAAGGCAC CACCCACCTA TGATACTGTG CTACAGATGG AGTATCTTGA CATGGTGGTG
1081 AATGAAACGC TCAGATTATT CCCAATTGCT ATGAGACTTG AGAGGGTCTG CAAAAAAGAT
1141 GTTGAGATCA ATGGGATGTT CATTCCCAAA GGGGTGGTGG TGATGATTCC AAGCTATGCT
1201 CTTCACCGTG ACCCAAAGTA CTGGACAGAG CCTGAGAAGT TCCTCCCTGA AAGATTCAGC
1261 AAGAAGAACA AGGACAACAT AGATCCTTAC ATATACACAC CCTTTGGAAG TGGACCCAGA
1321 AACTGCATTG GCATGAGGTT TGCTCTCATG AACATGAAAC TTGCTCTAAT CAGAGTCCTT
1381 CAGAACTTCT CCTTCAAACC TTGTAAAGAA ACACAGATCC CCCTGAAATT AAGCTTAGGA
1441 GGACTTCTTC AACCAGAAAA ACCCGTTGTT CTAAAGGTTG AGTCAAGGGA TGGCACCGTA
1501 AGTGGAGCCT GA
```

FIG. 11A

```
  1 MALIPDLAME TWLLLAVSLV LLYLYGTHSH GLFKKLGIPG PTPLPFLGNI LSYHKGFCMF
 61 DMECHKKYGK VWGFYDGQQP VLAITDPDMI KTVLVKECYS VFTNRRPFGP VGFMKSAISI
121 AEDEEWKRLR SLLSPTFTSG KLKEMVPIIA QYGDVLVRNL RREAETGKPV TLKDVFGAYS
181 MDVITSTSFG VNIDSLNNPQ DPFVENTKKL LRFDFLDPFF LSITVFPFLI PILEVLNICV
241 FPREVTNFLR KSVKRMKESR LEDTQKHRVD FLQLMIDSQN SKETESHKAL SDLELVAQSI
301 IFIFAGYETT SSVLSFIMYE LATHPDVQQK LQEEIDAVLP NKAPPTYDTV LQMEYLDMVV
361 NETLRLFPIA MRLERVCKKD VEINGMFIPK GVVVMIPSYA LHRDPKYWTE PEKFLPERFS
421 KKNKDNIDPY IYTPFGSGPR NCIGMRFALM NMKLALIRVL QNFSFKPCKE TQIPLKLSLG
481 GLLQPEKPVV LKVESRDGTV SGA*
```

FIG. 11B

```
   1 ATGGATTCTC TTGTGGTCCT TGTGCTCTGT CTCTCATGTT TGCTTCTCCT TTCACTCTGG
  61 AGACAGAGCT CTGGGAGAGG AAAACTCCCT CCTGGCCCCA CTCCTCTCCC AGTGATTGGA
 121 AATATCCTAC AGATAGGTAT TAAGGACATC AGCAAATCCT TAACCAATCT CTCAAAGGTC
 181 TATGGCCCGG TGTTCACTCT GTATTTTGGC CTGAAACCCA TAGTGGTGCT GCATGGATAT
 241 GAAGCAGTGA AGGAAGCCCT GATTGATCTT GGAGAGGAGT TTTCTGGAAG AGGCATTTTC
 301 CCACTGGCTG AAAGAGCTAA CAGAGGATTT GGAATTGTTT TCAGCAATGG AAAGAAATGG
 361 AAGGAGATCC GGCGTTTCTC CCTCATGACG CTGCGGAATT TTGGGATGGG GAAGAGGAGC
 421 ATTGAGGACC GTGTTCAAGA GGAAGCCCGC TGCCTTGTGG AGGAGTTGAG AAAAAACCAAG
 481 GCCTCACCCT GTGATCCCAC TTTCATCCTG GGCTGTGCTC CCTGCAATGT GATCTGCTCC
 541 ATTATTTTCC ATAAACGTTT TGATTATAAA GATCAGCAAT TTCTTAACTT AATGGAAAAG
 601 TTGAATGAAA ACATCAAGAT TTTGAGCAGC CCCTGGATCC AGATCTGCAA TAATTTTTCT
 661 CCTATCATTG ATTACTTCCC GGGAACTCAC AACAAATTAC TTAAAAACGT TGCTTTTATG
 721 AAAAGTTATA TTTTGGAAAA AGTAAAAGAA CACCAAGAAT CAATGGACAT GAACAACCCT
 781 CAGGACTTTA TTGATTGCTT CCTGATGAAA ATGGAGAAGG AAAAGCACAA CCAACCATCT
 841 GAATTTACTA TTGAAAGCTT GGAAAACACT GCAGTTGACT TGTTTGGAGC TGGGACAGAG
 901 ACGACAAGCA CAACCCTGAG ATATGCTCTC CTTCTCCTGC TGAAGCACCC AGAGGTCACA
 961 GCTAAAGTCC AGGAAGAGAT TGAACGTGTG ATTGGCAGAA ACCGGAGCCC CTGCATGCAA
1021 GACAGGAGCC ACATGCCCTA CACAGATGCT GTGGTGCACG AGGTCCAGAG ATACATTGAC
1081 CTTCTCCCCA CCAGCCTGCC CCATGCAGTG ACCTGTGACA TTAAATTCAG AAACTATCTC
1141 ATTCCCAAGG GCACAACCAT ATTAATTTCC CTGACTTCTG TGCTACATGA CAACAAAGAA
1201 TTTCCCAACC CAGAGATGTT TGACCCTCAT CACTTTCTGG ATGAAGGTGG CAATTTTAAG
1261 AAAAGTAAAT ACTTCATGCC TTTCTCAGCA GGAAAACGGA TTTGTGTGGG AGAAGCCCTG
1321 GCCGGCATGG AGCTGTTTTT ATTCCTGACC TCCATTTTAC AGAACTTTAA CCTGAAATCT
1381 CTGGTTGACC CAAAGAACCT TGACACCACT CCAGTTGTCA ATGGATTTGC CTCTGTGCCG
1441 CCCTTCTACC AGCTGTGCTT CATTCCTGTC TGAAGAAGAG CAGATGGCCT GGCTGCTGCT
1501 GTGCAGTCCC TGCAGCTCTC TTTCCTCTGG GGCATTATCC ATCTTTGCAC TATCTGTAAT
1561 GCCTTTTCTC ACCTGTCATC TCACATTTTC CCTTCCCTGA AGATCTAGTG AACATTCGAC
1621 CTCCATTACG GAGAGTTTCC TATGTTTCAC TGTGCAAATA TATCTGCTAT TCTCCATACT
1681 CTGTAACAGT TGCATTGACT GTCACATAAT GCTCATACTT ATCTAATGTA GAGTATTAAT
1741 ATGTTATTAT TAAATAGAGA AATATGATTT GTGTATTATA ATTCAAAGGC ATTTCTTTTC
1801 TGCATGATCT AAATAAAAAG CATTATTATT TGCTG
```

FIG. 12A

```
  1 MDSLVVLVLC LSCLLLLSLW RQSSGRGKLP PGPTPLPVIG NILQIGIKDI SKSLTNLSKV
 61 YGPVFTLYFG LKPIVVLHGY EAVKEALIDL GEEFSGRGIF PLAERANRGF GIVFSNGKKW
121 KEIRRFSLMT LRNFGMGKRS IEDRVQEEAR CLVEELRKTK ASPCDPTFIL GCAPCNVICS
181 IIFHKRFDYK DQQFLNLMEK LNENIKILSS PWIQICNNFS PIIDYFPGTH NKLLKNVAFM
241 KSYILEKVKE HQESMDMNNP QDFIDCFLMK MEKEKHNQPS EFTIESLENT AVDLFGAGTE
301 TTSTTLRYAL LLLLKHPEVT AKVQEEIERV IGRNRSPCMQ DRSHMPYTDA VVHEVQRYID
361 LLPTSLPHAV TCDIKFRNYL IPKGTTILIS LTSVLHDNKE FPNPEMFDPH HFLDEGGNFK
421 KSKYFMPFSA GKRICVGEAL AGMELFLFLT SILQNFNLKS LVDPKNLDTT PVVNGFASVP
481 PFYQLCFIPV *RRADGLAAA VQSLQLSFLW GIIHLCTICN AFSHLSSHIF PSLKI**TFD
541 LHYGEFPMFH CANISAILHT L*QLH*LSHN AHTYLM*SIN MLLLNREI*F VYYNSKAFLF
601 CMI*IKSIII C
```

FIG. 12B

```
   1 ATGGGGCTAG AAGCACTGGT GCCCCTGGCC GTGATAGTGG CCATCTTCCT GCTCCTGGTG
  61 GACCTGATGC ACCGGCGCCA ACGCTGGGCT GCACGCTACC CACCAGGCCC CCTGCCACTG
 121 CCCGGGCTGG GCAACCTGCT GCATGTGGAC TTCCAGAACA CACCATACTG CTTCGACCAG
 181 TTGCGGCGCC GCTTCGGGGA CGTGTTCAGC CTGCAGCTGG CCTGGACGCC GGTGGTCGTG
 241 CTCAATGGGC TGGCGGCCGT GCGCGAGGCG CTGGTGACCC ACGGCGAGGA CACCGCCGAC
 301 CGCCCGCCTG TGCCCATCAC CCAGATCCTG GGTTTCGGGC CGCGTTCCCA AGGGGTGTTC
 361 CTGGCGCGCT ATGGGCCCGC GTGGCGCGAG CAGAGGCGCT TCTCCGTGTC CACCTTGCGC
 421 AACTTGGGCC TGGGCAAGAA GTCGCTGGAG CAGTGGGTGA CCGAGGAGGC CGCCTGCCTT
 481 TGTGCCGCCT TCGCCAACCA CTCCGGACGC CCCTTTCGCC CCAACGGTCT CTTGGACAAA
 541 GCCGTGAGCA ACGTGATCGC CTCCCTCACC TGCGGGCGCC GCTTCGAGTA CGACGACCCT
 601 CGCTTCCTCA GGCTGCTGGA CCTAGCTCAG GAGGGACTGA AGGAGGAGTC GGGCTTTCTG
 661 CGCGAGGTGC TGAATGCTGT CCCCGTCCTC CTGCATATCC CAGCGCTGGC TGGCAAGGTC
 721 CTACGCTTCC AAAAGGCTTT CCTGACCCAG CTGGATGAGC TGCTAACTGA GCACAGGATG
 781 ACCTGGGACC CAGCCCAGCC CCCCCGAGAC CTGACTGAGG CCTTCCTGGC AGAGATGGAG
 841 AAGGCCAAGG GGAACCCTGA GAGCAGCTTC AATGATGAGA ACCTGCGCAT AGTGGTGGCT
 901 GACCTGTTCT CTGCCGGGAT GGTGACCACC TCGACCACGC TGGCCTGGGG CCTCCTGCTC
 961 ATGATCCTAC ATCCGGATGT GCAGCGCCGT GTCCAACAGG AGATCGACGA CGTGATAGGG
1021 CAGGTGCGGC GACCAGAGAT GGGTGACCAG GCTCACATGC CCTACACCAC TGCCGTGATT
1081 CATGAGGTGC AGCGCTTTGG GGACATCGTC CCCCTGGGTA TGACCCATAT GACATCCCGT
1141 GACATCGAAG TACAGGGCTT CCGCATCCCT AAGGGAACGA CACTCATCAC CAACCTGTCA
1201 TCGGTGCTGA AGGATGAGGC CGTCTGGGAG AAGCCCTTCC GCTTCCACCC CGAACACTTC
1261 CTGGATGCCC AGGGCCACTT TGTGAAGCCG GAGGCCTTCC TGCCTTTCTC AGCAGGCCGC
1321 CGTGCATGCC TCGGGGAGCC CCTGGCCCGC ATGGAGCTCT TCCTCTTCTT CACCTCCCTG
1381 CTGCAGCACT TCAGCTTCTC GGTGCCCACT GGACAGCCCC GGCCCAGCCA CCATGGTGTC
1441 TTTGCTTTCC TGGTGAGCCC ATCCCCCTAT GAGCTTTGTG CTGTGCCCCG CTAG
```

FIG. 13A

```
  1 MGLEALVPLA VIVAIFLLLV DLMHRRQRWA ARYPPGPLPL PGLGNLLHVD FQNTPYCFDQ
 61 LRRRFGDVFS LQLAWTPVVV LNGLAAVREA LVTHGEDTAD RPPVPITQIL GFGPRSQGVF
121 LARYGPAWRE QRRFSVSTLR NLGLGKKSLE QWVTEEAACL CAAFANHSGR PFRPNGLLDK
181 AVSNVIASLT CGRRFEYDDP RFLRLLDLAQ EGLKEESGFL REVLNAVPVL LHIPALAGKV
241 LRFQKAFLTQ LDELLTEHRM TWDPAQPPRD LTEAFLAEME KAKGNPESSF NDENLRIVVA
301 DLFSAGMVTT STTLAWGLLL MILHPDVQRR VQQEIDDVIG QVRRPEMGDQ AHMPYTTAVI
361 HEVQRFGDIV PLGMTHMTSR DIEVQGFRIP KGTTLITNLS SVLKDEAVWE KPFRFHPEHF
421 LDAQGHFVKP EAFLPFSAGR RACLGEPLAR MELFLFFTSL LQHFSFSVPT GQPRPSHHGV
481 FAFLVSPSPY ELCAVPR*
```

FIG. 13B

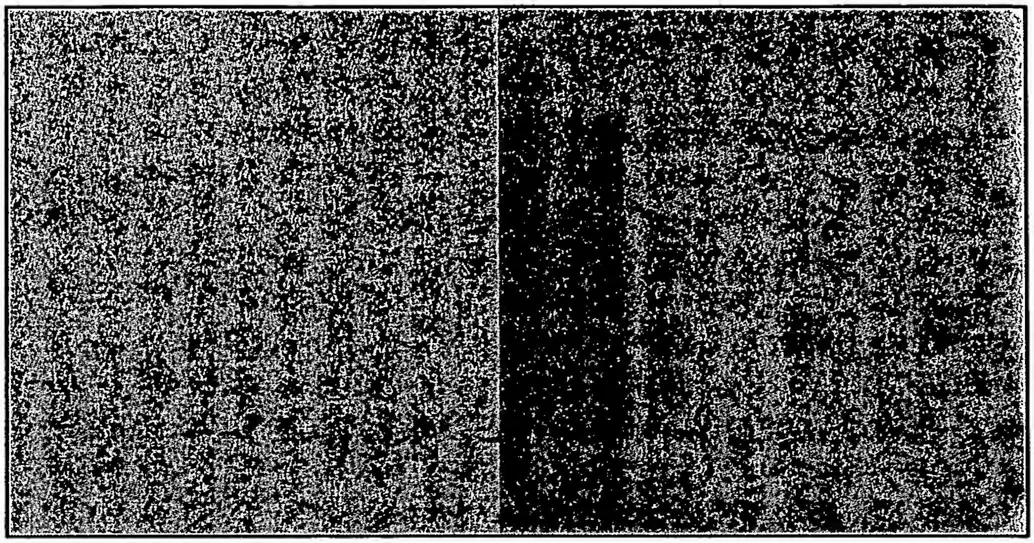


FIG. 14

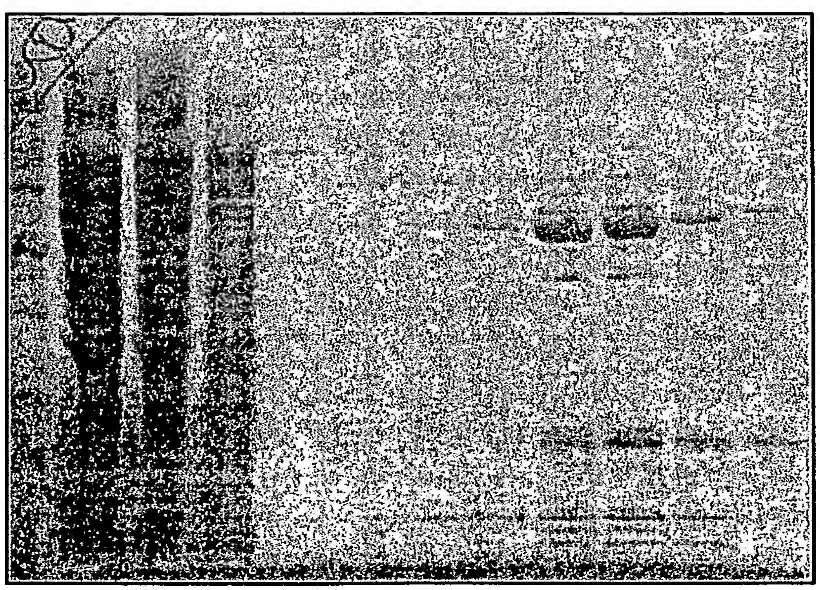


FIG. 15

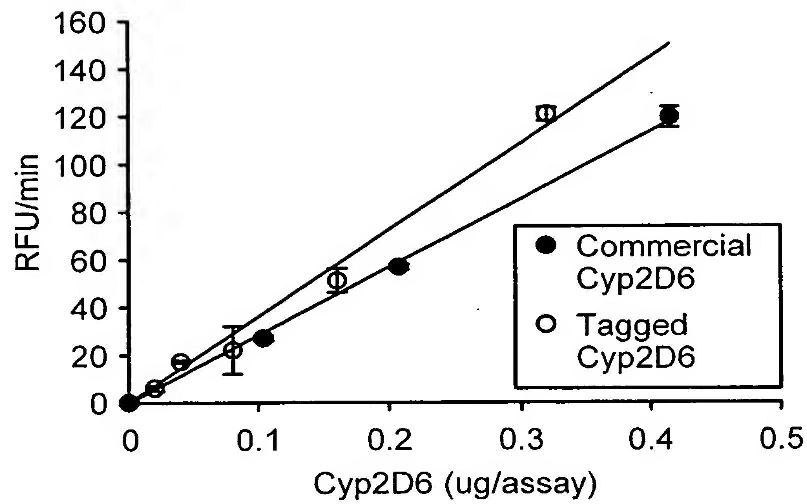

160
140
120
100
80
60
40
20
0
RFU/min
0
0.1
0.2
0.3
0.4
0.5
Cyp2D6 (ug/assay)
Commercial Cyp2D6
Tagged Cyp2D6

FIG. 16

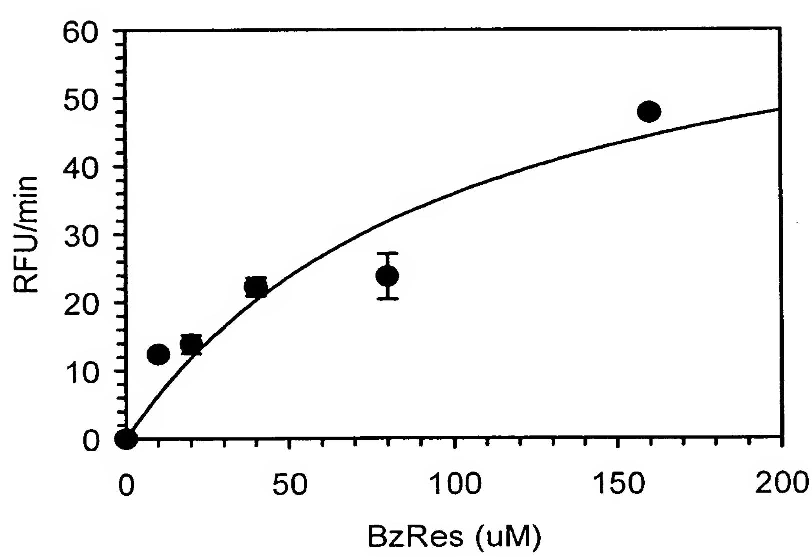

60
50
40
30
20
10
0
RFU/min
0
50
100
150
200
BzRes (uM)

FIG. 17

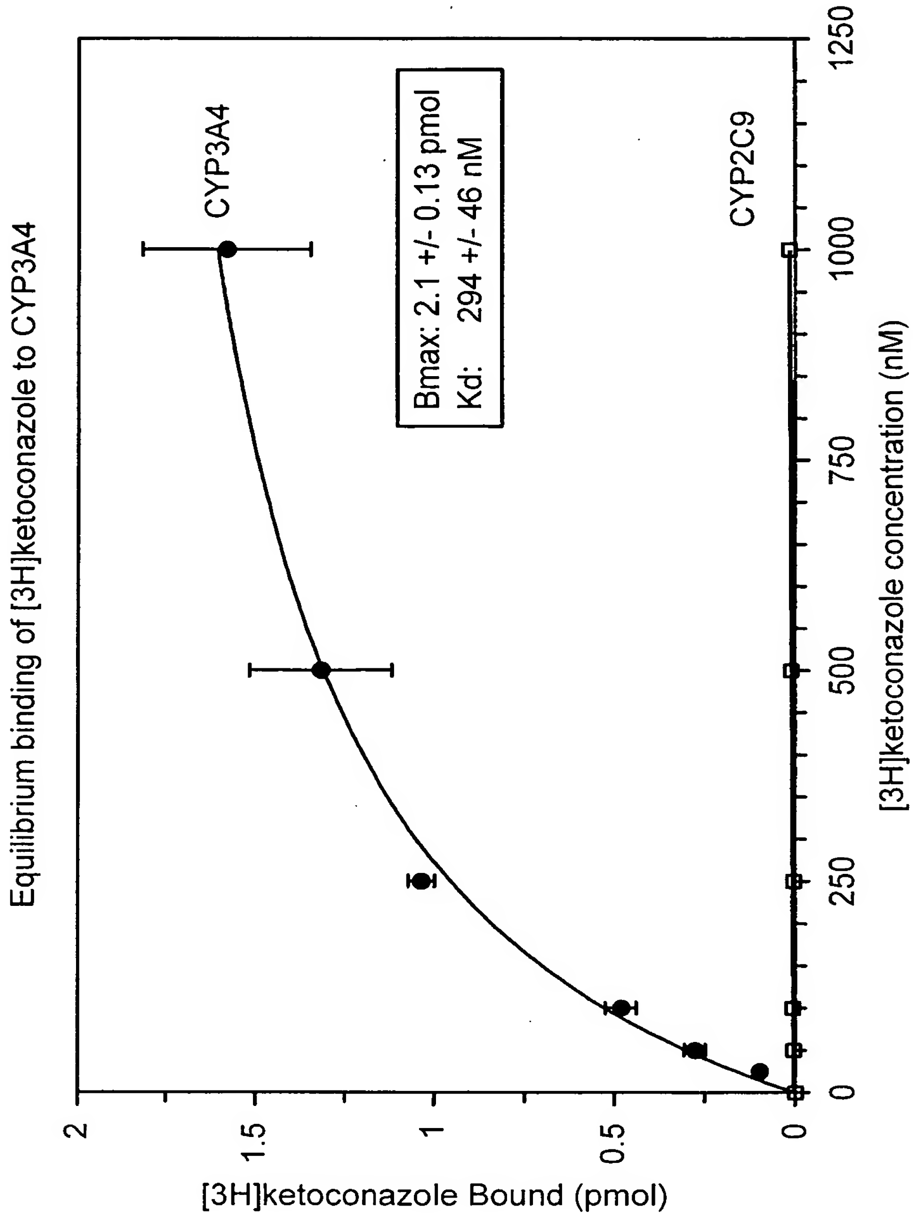

Equilibrium binding of [3H]ketoconazole to CYP3A4
CYP3A4
CYP2C9
Bmax: 2.1 +/- 0.13 pmol
Kd: 294 +/- 46 nM
[3H]ketoconazole Bound (pmol)
0
0.5
1
1.5
2
[3H]ketoconazole concentration (nM)
0
250
500
750
1000
1250

FIG. 18

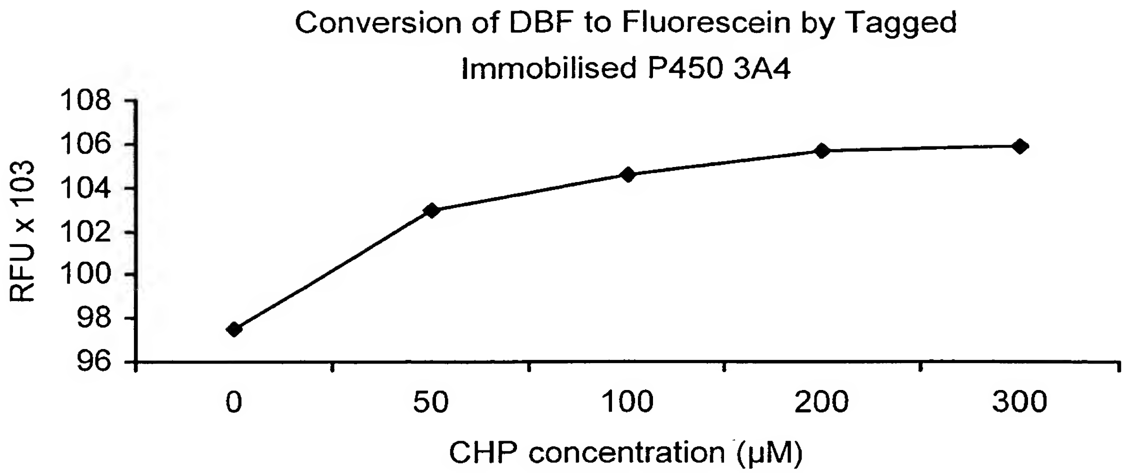

Conversion of DBF to Fluorescein by Tagged Immobilised P450 3A4
RFU x 103
108
106
104
102
100
98
96
0
50
100
200
300
CHP concentration (μM)

FIG. 19

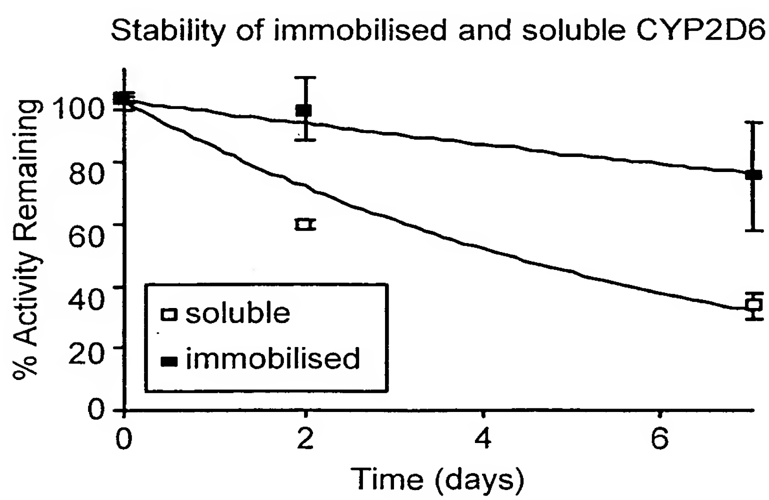

Stability of immobilised and soluble CYP2D6
% Activity Remaining
100
80
60
40
20
0
soluble
immobilised
0
2
4
6
Time (days)

FIG. 20

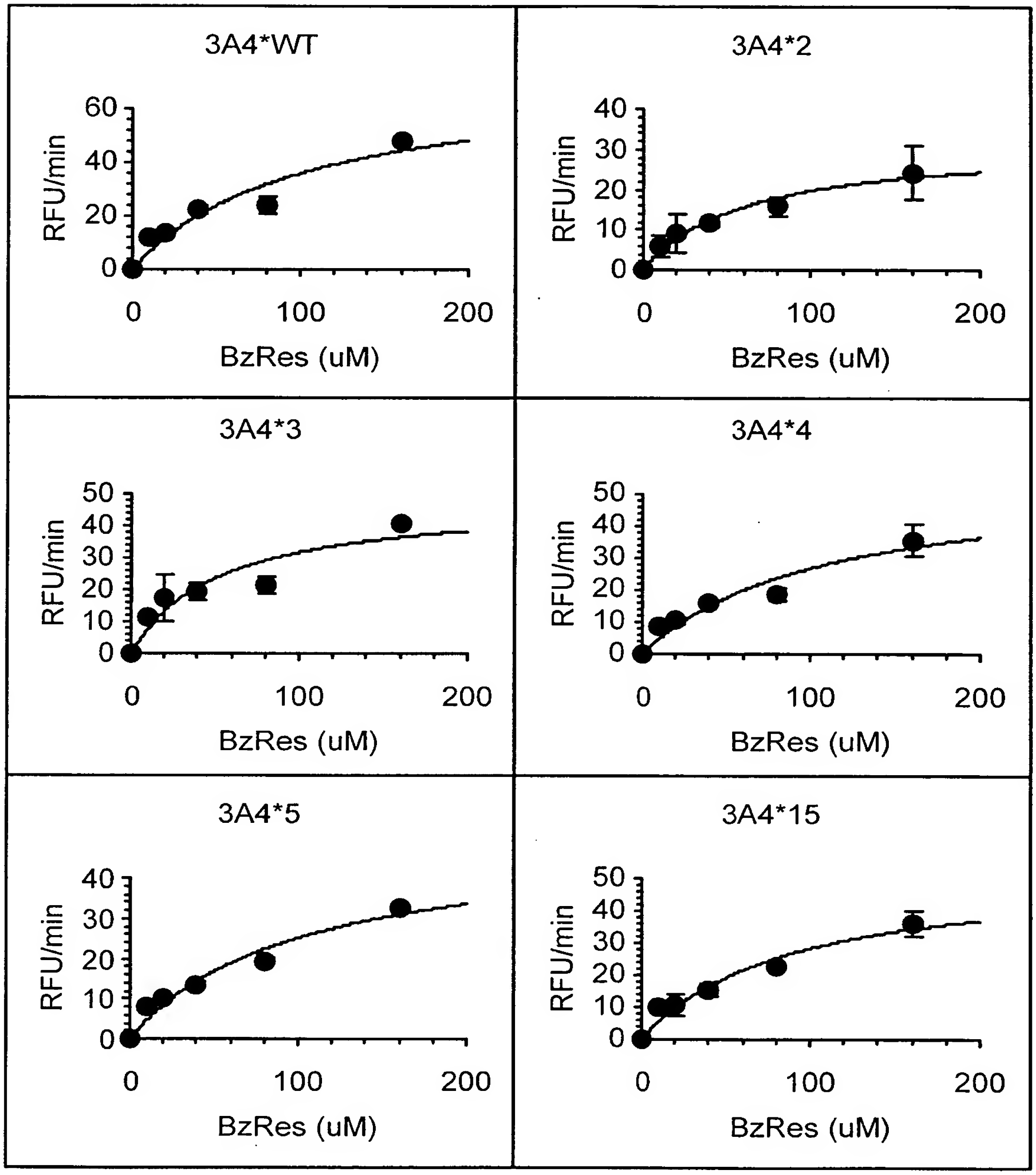

3A4*WT
3A4*2
3A4*3
3A4*4
3A4*5
3A4*15
RFU/min
BzRes (uM)
0
100
200
0
20
40
60
10
30
50

FIG. 21

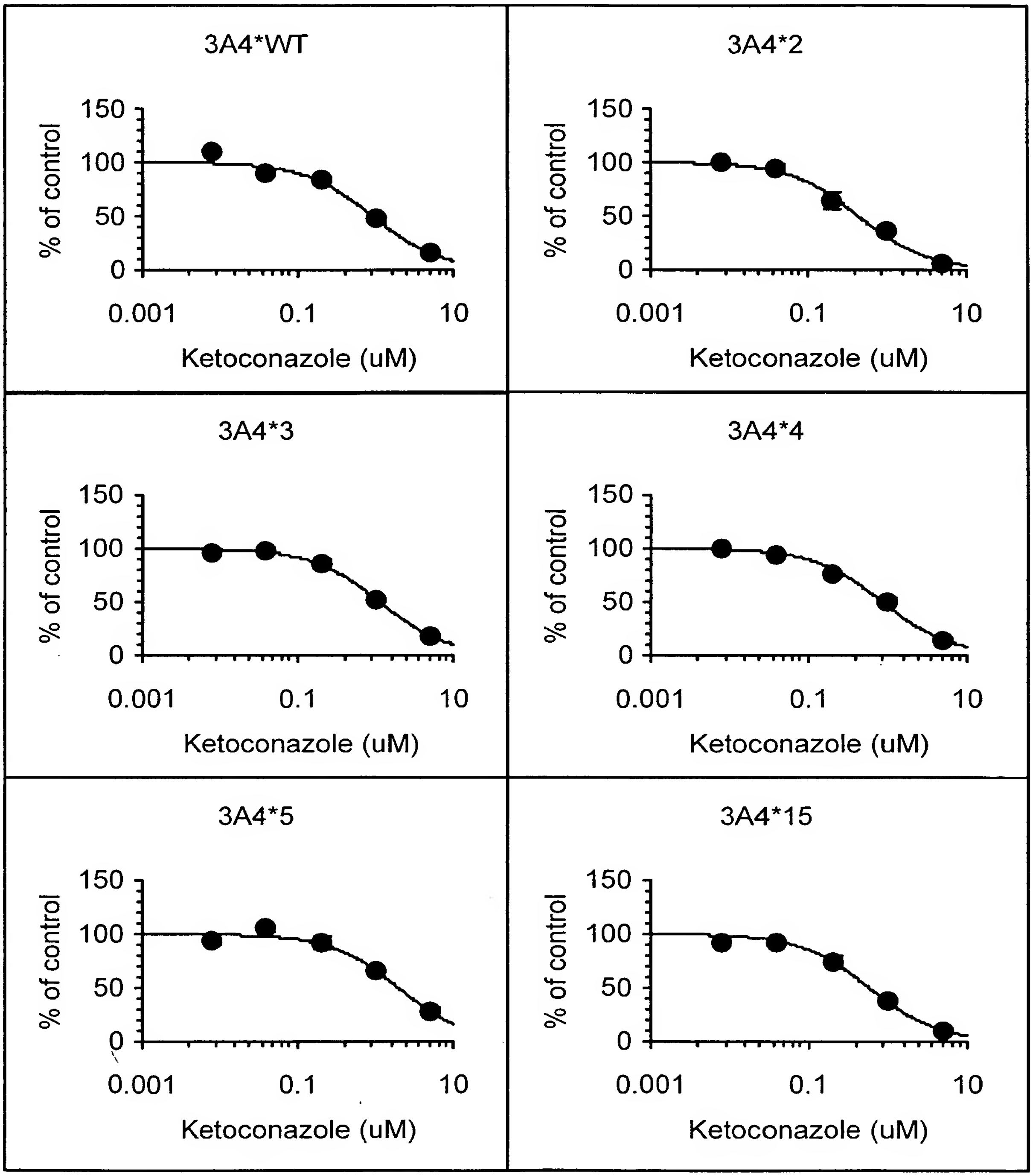

3A4*WT
3A4*2
3A4*3
3A4*4
3A4*5
3A4*15
% of control
150
100
50
0
0.001
0.1
10
Ketoconazole (uM)

FIG. 22